Estimation of Camera Pose, Linear

Computer Vision II

CSE 252B

Lecture 7

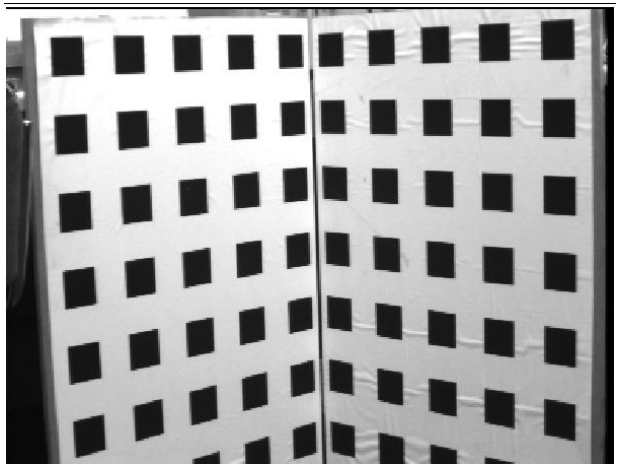
Announcements

- Assignment 2 is due Feb 4, 11:59 PM
- Reading
  - Lepetit, Moreno-Noguer, and Fua, “EPnP: An Accurate O(n) Solution to the PnP Problem”

Calibrated camera

Camera calibration

- Use calibration procedure to determine camera calibration matrix (and lens distortion parameters)
  - One or more images of a calibration object

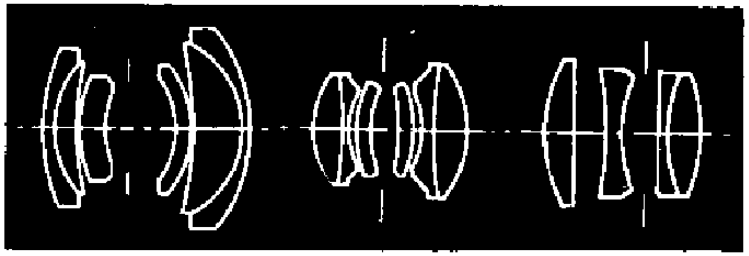


Example image of calibration object

3D points on calibration object are known

Lens distortion

- Radial distortion
- Tangential distortion



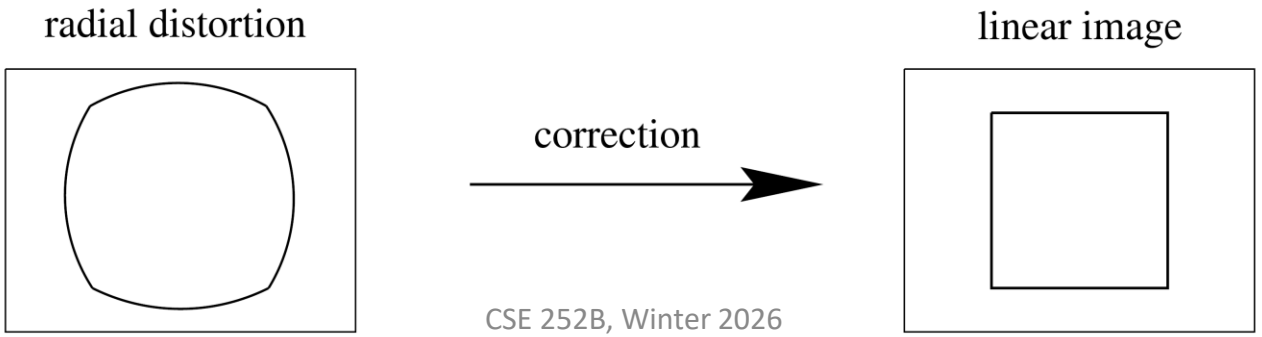


Example distorted image

Images of straight lines should be straight



Example undistorted image



Camera calibration matrix

$$\mathtt{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where

$\alpha_x = f m_x$	Focal length in x direction in terms of pixel dimensions
$\alpha_y = f m_y$	Focal length in y direction in terms of pixel dimensions
s	Skew
$(x_0, y_0)^\top = (p_x m_x, p_y m_y)^\top$	Coordinates of principal point in terms of pixel dimensions

where m_x and m_y are number of pixels per unit disance in x and y directions, respectively

(and lens distortion parameters)

Single view geometry

Camera projection

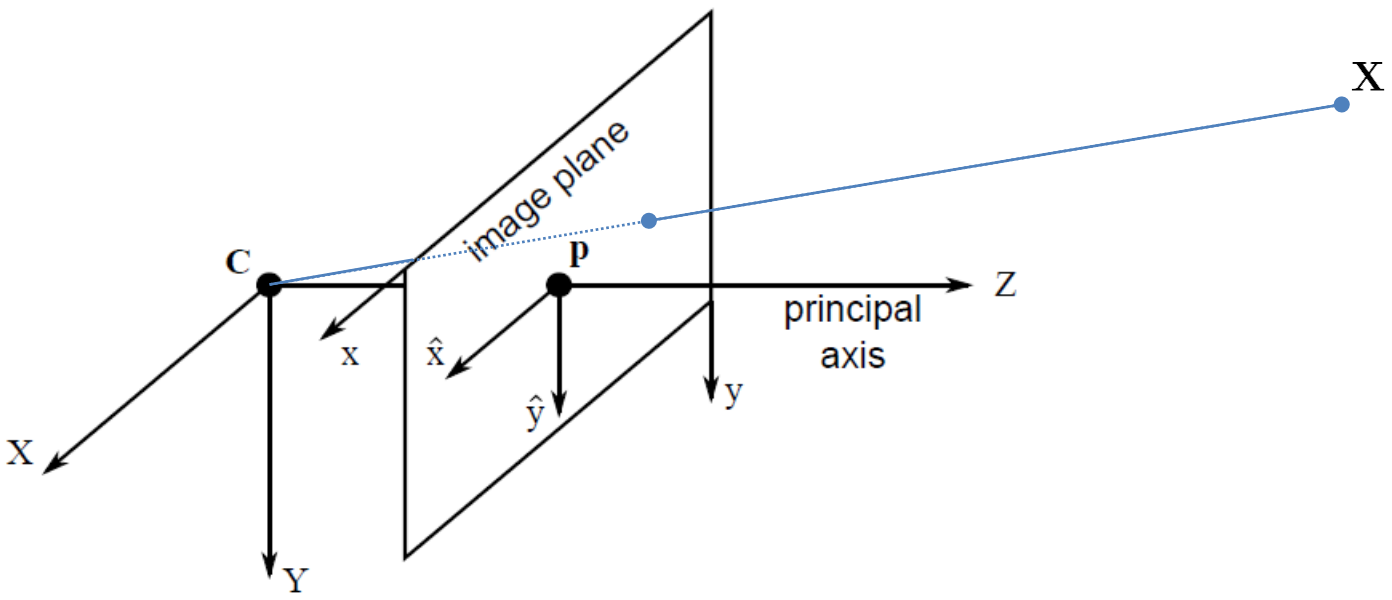
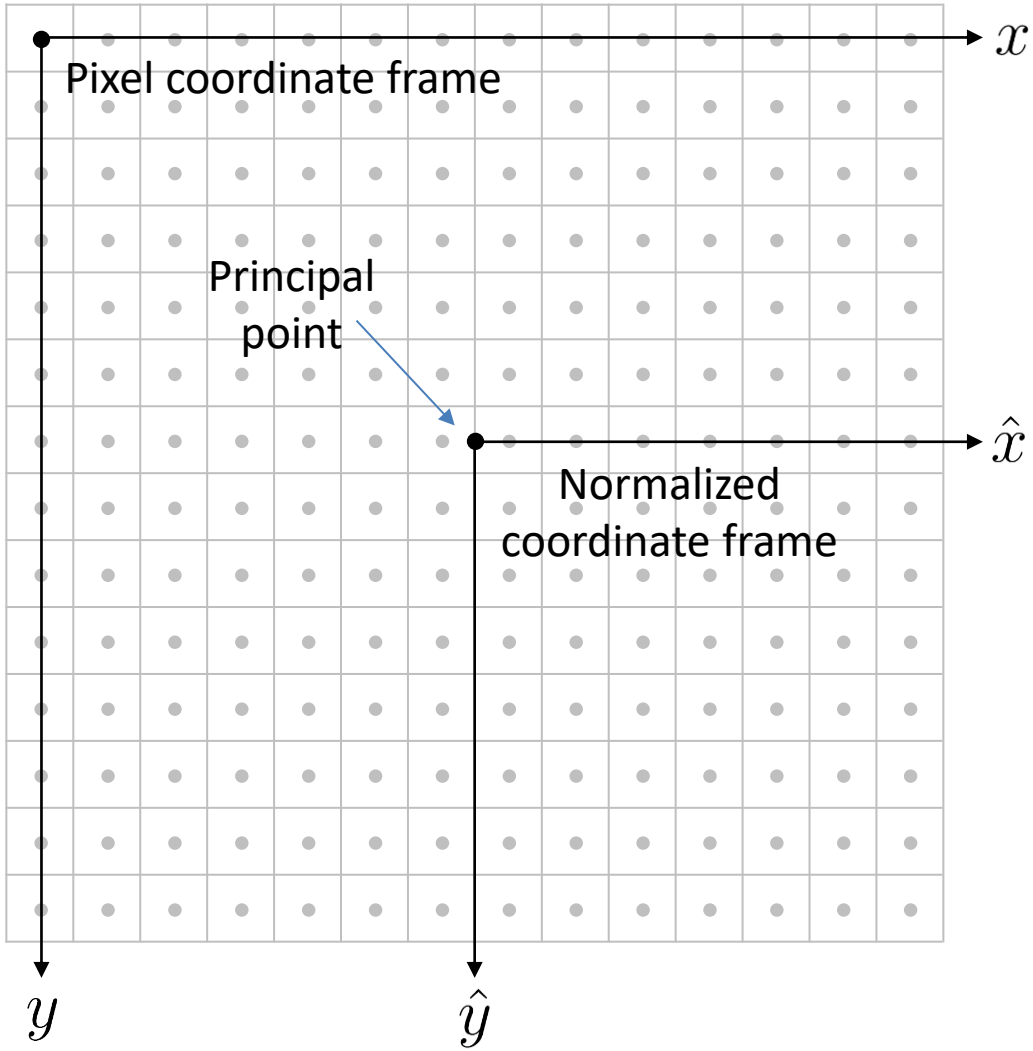


Image coordinate frames



Camera projection

Project 3D point in world coordinate frame $\mathbf{X}$ under camera projection matrix $\mathtt{P} = \mathtt{K}[\mathtt{R} \,|\, \mathbf{t}]$

$$\begin{aligned} \mathbf{x} &= \mathtt{P}\mathbf{X} \\ \mathbf{x} &= \mathtt{K}[\mathtt{R} \,|\, \mathbf{t}]\mathbf{X} \\ \mathtt{K}^{-1}\mathbf{x} &= [\mathtt{R} \,|\, \mathbf{t}]\mathbf{X} \\ \hat{\mathbf{x}} &= \hat{\mathtt{P}}\mathbf{X} \end{aligned}$$

Camera calibration matrix

where

Camera projection matrix	$\mathtt{P} = \mathtt{K}[\mathtt{R} \,|\, \mathbf{t}]$	$= \mathtt{K}\hat{\mathtt{P}}$
Normalized camera projection matrix	$\hat{\mathtt{P}} = [\mathtt{R} \,|\, \mathbf{t}]$	$= \mathtt{K}^{-1}\mathtt{P}$
Image point in pixel coordinates	$\mathbf{x} = \mathtt{K}[\mathtt{R} \,|\, \mathbf{t}]\mathbf{X}$	$= \mathtt{K}\hat{\mathbf{x}}$
Image point in normalized coordinates	$\hat{\mathbf{x}} = [\mathtt{R} \,|\, \mathbf{t}]\mathbf{X}$	$= \mathtt{K}^{-1}\mathbf{x}$

Together, the camera rotation and camera translation (or camera center $\tilde{\mathbf{C}} = -\mathtt{R}^{\top}\mathbf{t}$) are called the camera pose. The normalized camera projection matrix $\hat{\mathtt{P}} = [\mathtt{R} \,|\, \mathbf{t}]$ embodies the camera pose.

Calibrated camera

$$\mathtt{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where

$\alpha_x = f m_x$	Focal length in x direction in terms of pixel dimensions
$\alpha_y = f m_y$	Focal length in y direction in terms of pixel dimensions
s	Skew
$(x_0, y_0)^\top = (p_x m_x, p_y m_y)^\top$	Coordinates of principal point in terms of pixel dimensions

where m_x and m_y are number of pixels per unit disance in x and y directions, respectively

- If camera calibration parameters are known, then use normalized camera projection matrix and image points in normalized coordinates

Transformation from world coordinate frame to camera coordinate frame

3D point

$$\tilde{\mathbf{X}}_{\text{cam}} = \mathtt{R}\tilde{\mathbf{X}} + \mathbf{t}$$

$$\begin{bmatrix} \tilde{\mathbf{X}}_{\text{cam}} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}} \\ 1 \end{bmatrix}$$

$$\mathbf{X}_{\text{cam}} = \mathtt{H}_{\text{world,cam}}\mathbf{X}, \text{ where } \mathtt{H}_{\text{world,cam}} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

Normalized camera

$$\hat{\mathbf{x}} = \hat{\mathtt{P}}\mathbf{X}$$

$$\hat{\mathbf{x}} = \hat{\mathtt{P}}\mathtt{H}^{-1}_{\text{world,cam}}\mathtt{H}_{\text{world,cam}}\mathbf{X}$$

$$\hat{\mathbf{x}} = \hat{\mathtt{P}}\mathtt{H}^{-1}_{\text{world,cam}}\mathbf{X}_{\text{cam}}, \text{ where } \mathbf{X}_{\text{cam}} = \mathtt{H}_{\text{world,cam}}\mathbf{X}$$

$$\hat{\mathbf{x}} = [\mathtt{R} \,|\, \mathbf{t}] \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}^{-1} \mathbf{X}_{\text{cam}}$$

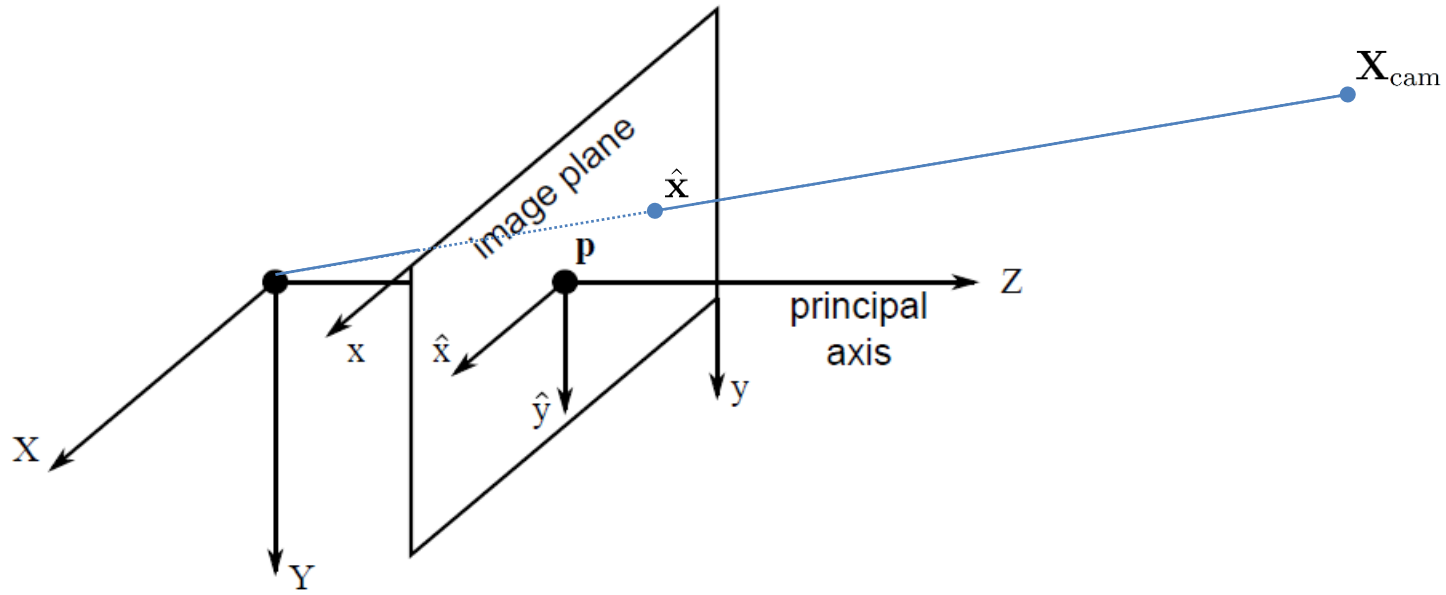
$$\hat{\mathbf{x}} = [\mathtt{I} \,|\, \mathbf{0}] \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}^{-1} \mathbf{X}_{\text{cam}}$$

$$\hat{\mathbf{x}} = [\mathtt{I} \,|\, \mathbf{0}]\mathbf{X}_{\text{cam}}, \text{ where } [\mathtt{I} \,|\, \mathbf{0}] = \hat{\mathtt{P}}\mathtt{H}^{-1}_{\text{world,cam}}$$

Camera projection

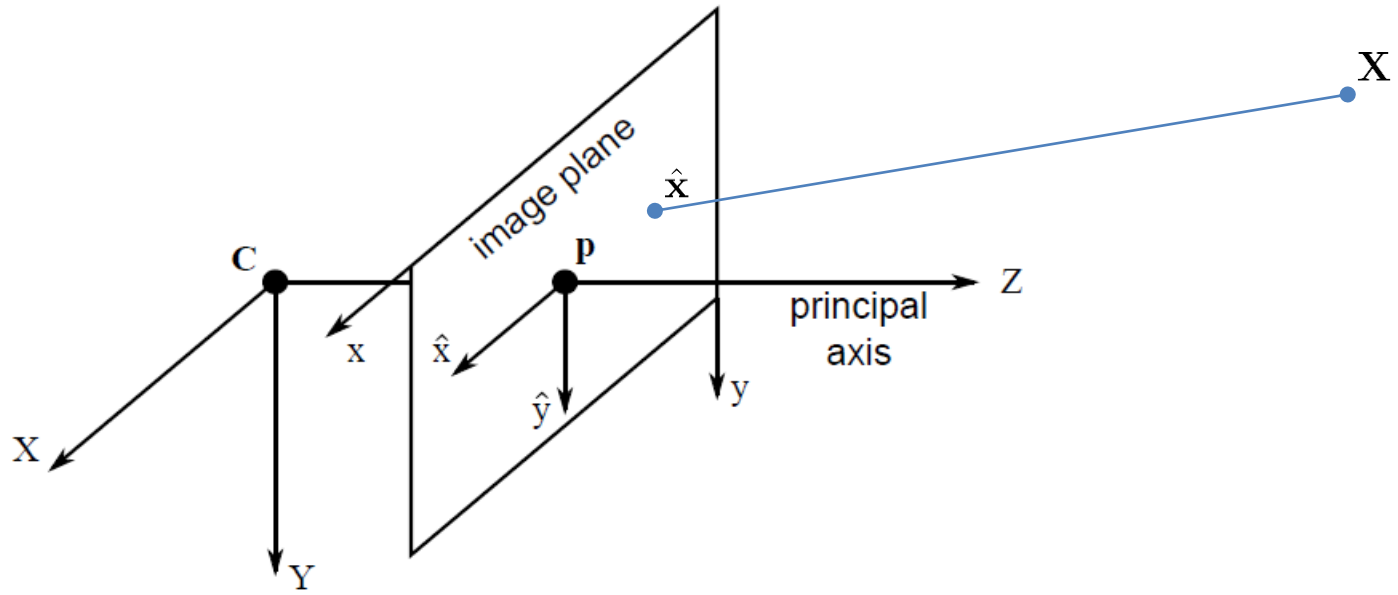
Transform points from world coordinate frame to camera coordinate frame

$$\tilde{\mathbf{X}}_{\text{cam}} = \mathtt{R}\tilde{\mathbf{X}} + \mathbf{t}$$

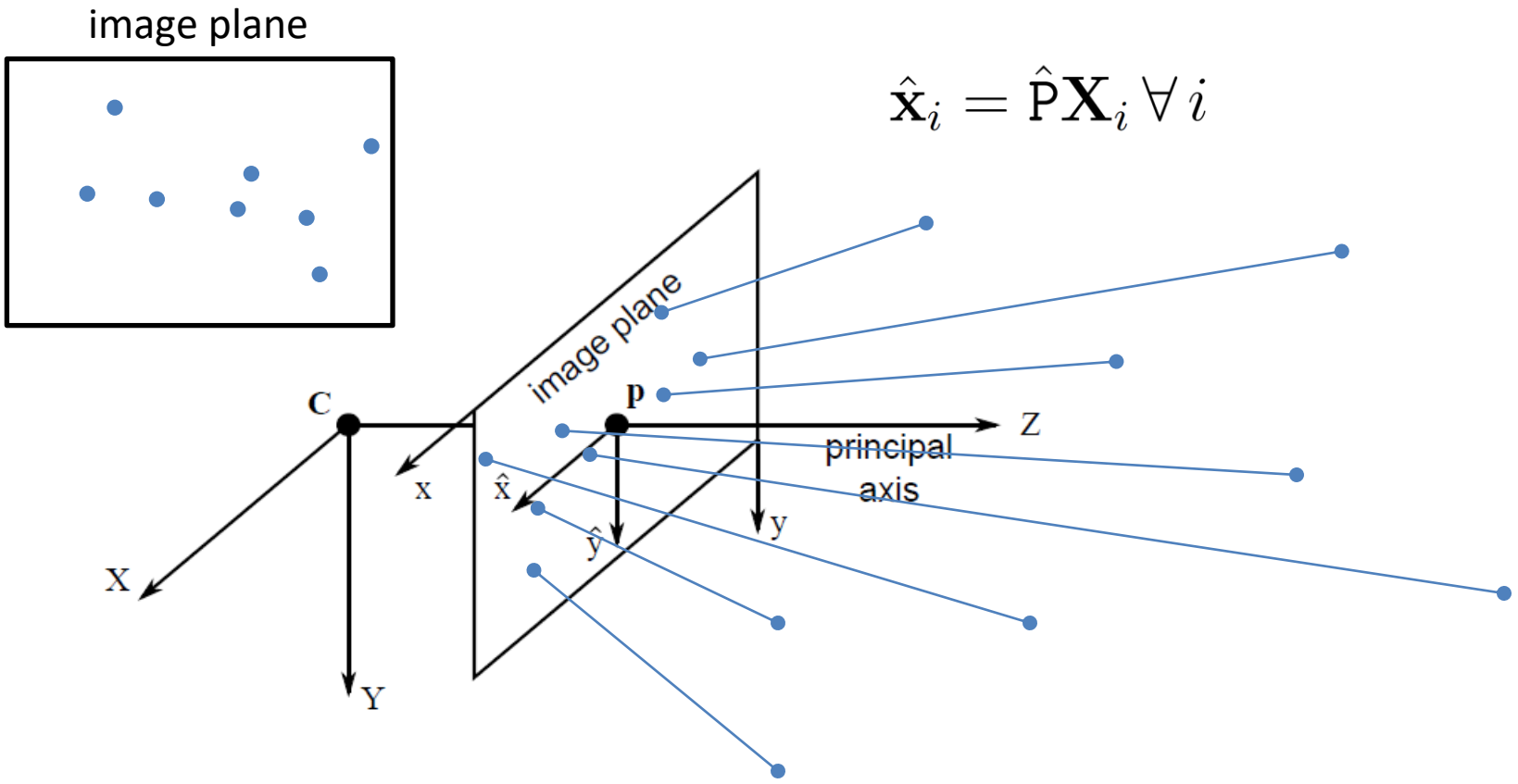


Normalized camera projection matrix

$$\hat{\mathbf{x}} = \hat{\mathrm{P}}\mathbf{X}$$

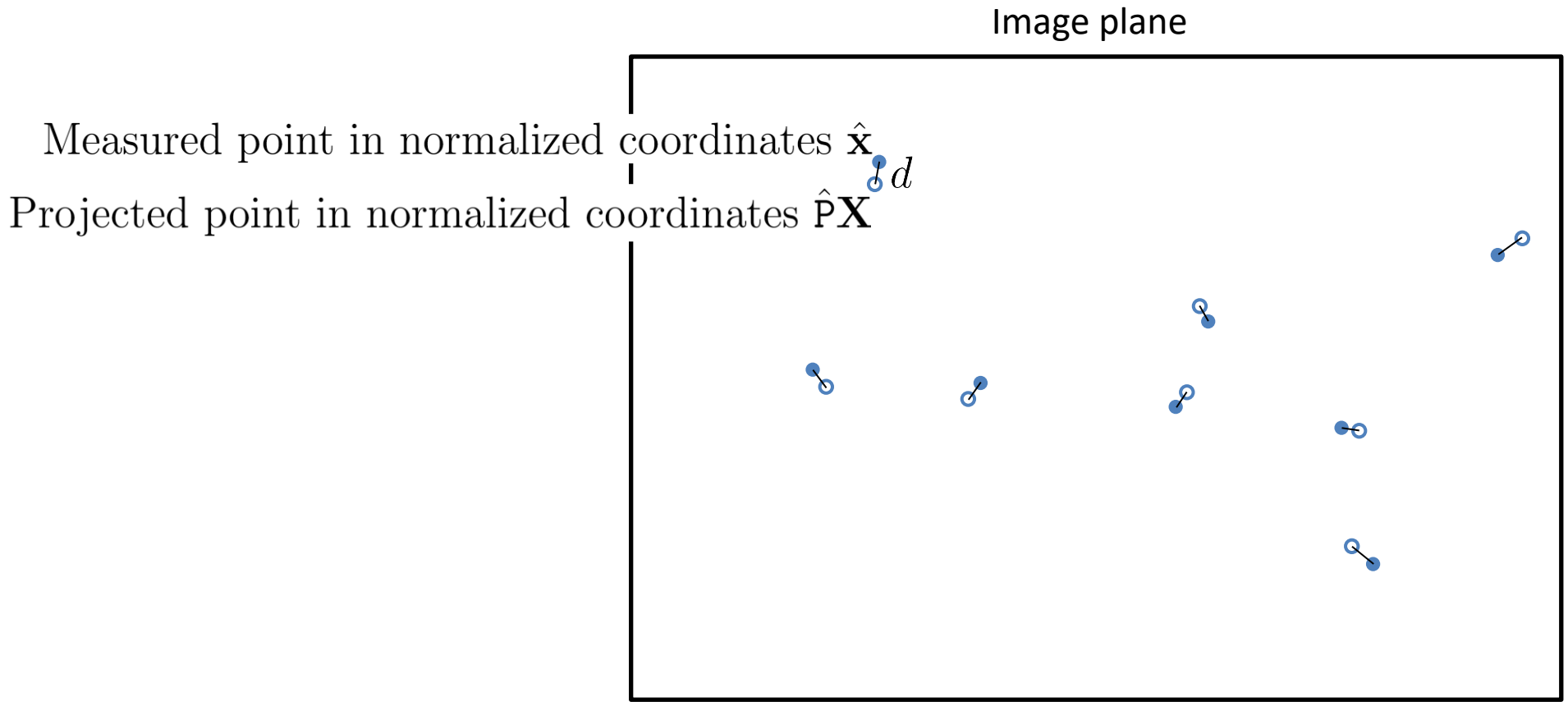


Normalized camera projection matrix



Camera projection matrix

- Error in image point measurements



Normalized camera projection matrix

Linear projection of points in homogeneous coordinates

$$\hat{\mathbf{x}} = \hat{\mathtt{P}}\mathbf{X}$$
$$\hat{\mathbf{x}} = [\mathtt{R} \,|\, \mathbf{t}]\mathbf{X}$$
$$\hat{\mathbf{x}} = [\mathtt{I} \,|\, \mathbf{0}]\mathbf{X}_{\text{cam}}$$
$$\hat{\mathbf{x}} = [\mathtt{I} \,|\, \mathbf{0}] \begin{bmatrix} \tilde{\mathbf{X}}_{\text{cam}} \\ 1 \end{bmatrix}$$
$$\hat{\mathbf{x}} = \tilde{\mathbf{X}}_{\text{cam}}, \text{ where } \tilde{\mathbf{X}}_{\text{cam}} = \mathtt{R}\tilde{\mathbf{X}} + \mathbf{t}$$
$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{w} \end{bmatrix} = \begin{bmatrix} \tilde{X}_{\text{cam}} \\ \tilde{Y}_{\text{cam}} \\ \tilde{Z}_{\text{cam}} \end{bmatrix}$$

Nonlinear projection of points in inhomogeneous coordinates

$$\hat{\tilde{x}} = \frac{\hat{x}}{\hat{w}} = \frac{\tilde{X}_{\text{cam}}}{\tilde{Z}_{\text{cam}}} \text{ and}$$
$$\hat{\tilde{y}} = \frac{\hat{y}}{\hat{w}} = \frac{\tilde{Y}_{\text{cam}}}{\tilde{Z}_{\text{cam}}}$$

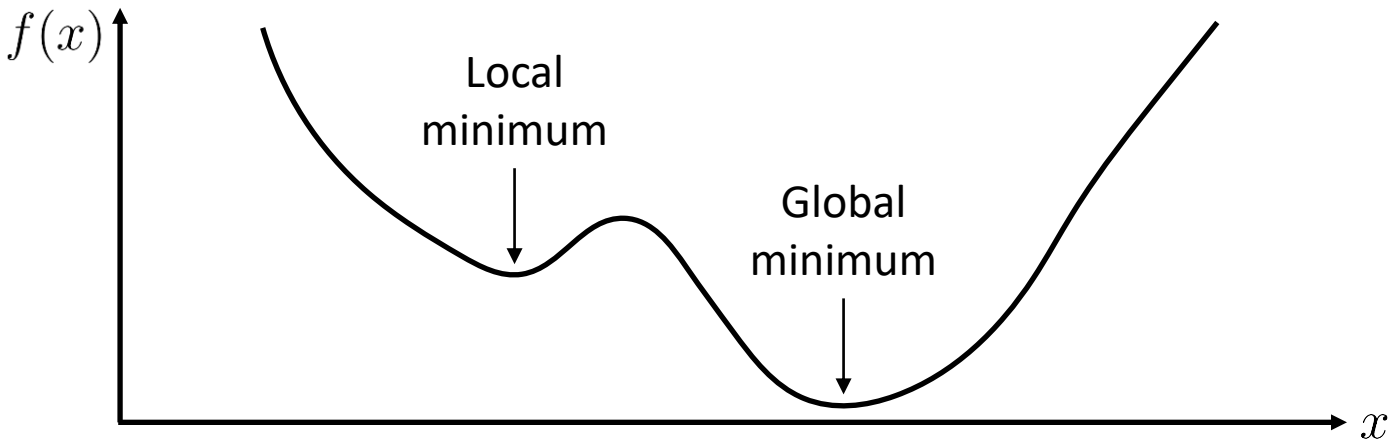
Do not forget this is a nonlinear projection

Estimation of normalized camera projection matrix

$$\hat{\mathbf{x}}_i = \hat{\mathtt{P}}\mathbf{X}_i \,\forall\, i$$

$$\hat{\mathbf{x}}_i = [\mathtt{R} \,|\, \mathbf{t}]\mathbf{X}_i \,\forall\, i, \text{ where } \hat{\mathbf{x}}_i = \mathtt{K}^{-1}\mathbf{x}_i \,\forall\, i \text{ and } \hat{\mathtt{P}} = [\mathtt{R} \,|\, \mathbf{t}], \text{ solve for } \mathtt{R} \text{ and } \mathbf{t}$$

- Given point correspondences $\hat{\mathbf{x}}_i \leftrightarrow \mathbf{X}_i$
- Minimize geometric error $\sum_i d(\hat{\mathbf{x}}_i, \hat{\mathtt{P}}\mathbf{X}_i)^2$
- A nonlinear mapping requires a nonlinear optimization problem solver
  - Use linear estimation for initial estimate (to mitigate converging to local optimum)
  - Iterative process to determine global optimum



Linear estimation of normalized camera projection matrix

Given point correspondences $\hat{\mathbf{x}}_i \leftrightarrow \tilde{\mathbf{X}}_i$

$$\hat{\mathbf{x}}_i = \hat{\mathtt{P}} \begin{bmatrix} \tilde{\mathbf{X}}_i \\ 1 \end{bmatrix} \forall i, \text{ solve for } \hat{\mathtt{P}}$$

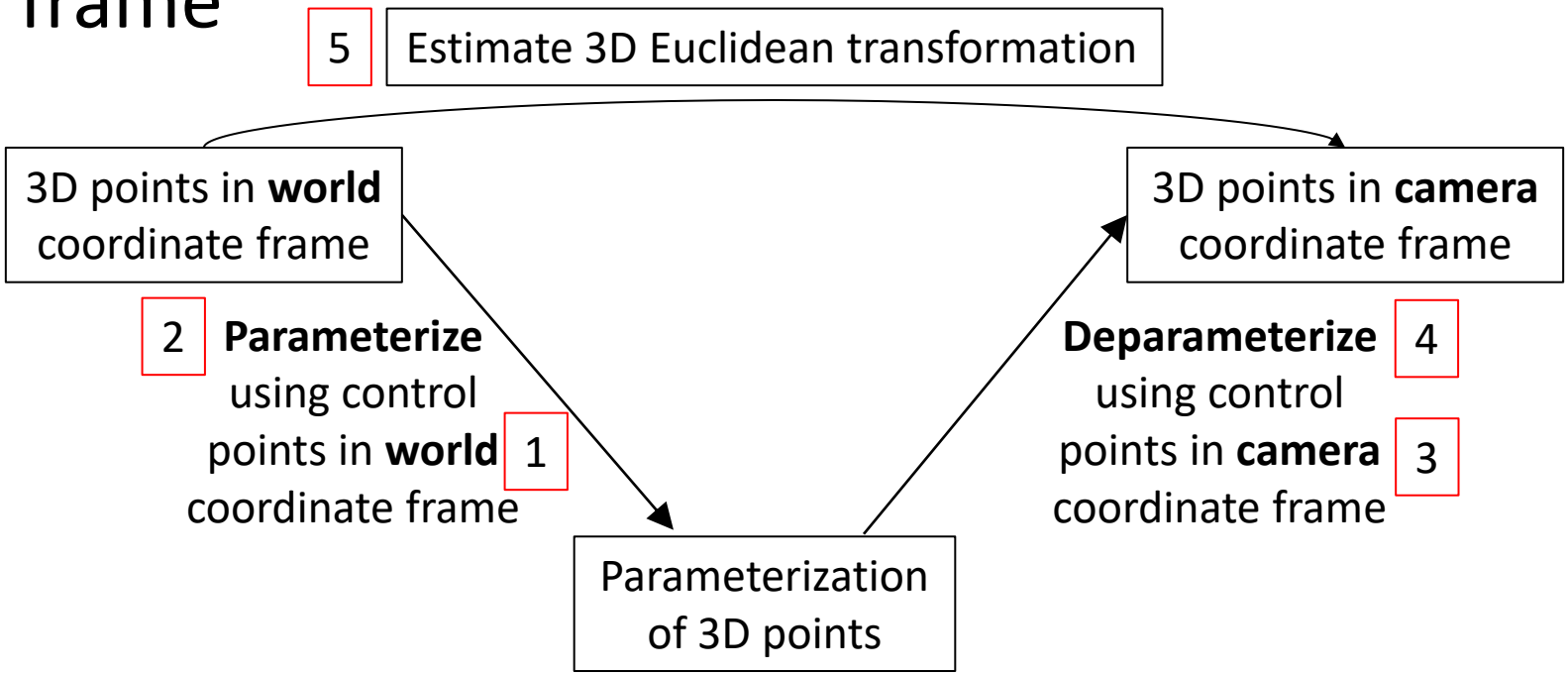
$$\hat{\mathbf{x}}_i = [\mathtt{R} \,|\, \mathbf{t}] \begin{bmatrix} \tilde{\mathbf{X}}_i \\ 1 \end{bmatrix} \forall i, \text{ where } \hat{\mathtt{P}} = [\mathtt{R} \,|\, \mathbf{t}], \text{ solve for } \mathtt{R} \text{ and } \mathbf{t}$$

$$\hat{\mathbf{x}}_i = \tilde{\mathbf{X}}_{\text{cam},i} \forall i, \text{ where } \tilde{\mathbf{X}}_{\text{cam},i} = \mathtt{R}\tilde{\mathbf{X}}_i + \mathbf{t} \forall i, \text{ solve for } \mathtt{R} \text{ and } \mathbf{t}$$

Determine set of 3D points in the camera coordinate frame corresponding to set of 3D points in the world coordinate frame (i.e., $\hat{\mathbf{X}}_{\text{cam},i} \leftrightarrow \tilde{\mathbf{X}}_i$, where $\tilde{\mathbf{X}}_{\text{cam},i} = \mathtt{R}\tilde{\mathbf{X}}_i + \mathbf{t} \forall i$), then solve for $\mathtt{R}$ and $\mathbf{t}$ using $\hat{\mathbf{X}}_{\text{cam},i} \leftrightarrow \tilde{\mathbf{X}}_i$

Linear estimation of normalized camera projection matrix

- Key to algorithm is parameterization of 3D points using control points in each coordinate frame



Linear estimation of normalized camera projection matrix

- Control points in world coordinate frame
  - From 3D points in world coordinate frame

Calculate the mean vector $\boldsymbol{\mu}_{\tilde{\mathbf{X}}}$ (i.e., centroid) and the covariance matrix $\Sigma_{\tilde{\mathbf{X}}}$ of the set of 3D points in world coordinate frame. Then, compute the eigen (or singular value) decomposition of the covariance matrix

$$\Sigma_{\tilde{\mathbf{X}}} = \mathtt{V}\Lambda\mathtt{V}^\top, \text{ where } \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3), \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0, \text{ and } \mathtt{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$$

$r = \text{rank}(\Sigma_{\tilde{\mathbf{X}}}) = \text{rank}(\Lambda)$. If $r = 2$, then 3D points are coplanar; if $r = 3$, then 3D points are in general position.

Coplanar, 3 control points

$$\tilde{\mathbf{C}}_1 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}}$$
$$\tilde{\mathbf{C}}_2 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}} + \mathbf{v}_1$$
$$\tilde{\mathbf{C}}_3 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}} + \mathbf{v}_2$$

General position, 4 control points

$$\tilde{\mathbf{C}}_1 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}}$$
$$\tilde{\mathbf{C}}_2 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}} + \mathbf{v}_1$$
$$\tilde{\mathbf{C}}_3 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}} + \mathbf{v}_2$$
$$\tilde{\mathbf{C}}_4 = \boldsymbol{\mu}_{\tilde{\mathbf{X}}} + \mathbf{v}_3$$

Parameterization of 3D points, coplanar

Given control points $\tilde{\mathbf{C}}_1, \tilde{\mathbf{C}}_2,$ and $\tilde{\mathbf{C}}_3$

$$\tilde{\mathbf{X}}_i = \sum_{j=1}^{3} \alpha_{i,j} \tilde{\mathbf{C}}_j \; \forall \, i$$

$$\tilde{\mathbf{X}}_i = \alpha_{i,1}\tilde{\mathbf{C}}_1 + \alpha_{i,2}\tilde{\mathbf{C}}_2 + \alpha_{i,3}\tilde{\mathbf{C}}_3, \text{ where } \alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} = 1 \, \forall \, i$$

solve for parameterization $(\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3})$ of i-th point $\tilde{\mathbf{X}}_i$

$$\tilde{\mathbf{X}}_i = (1 - \alpha_{i,2} - \alpha_{i,3})\tilde{\mathbf{C}}_1 + \alpha_{i,2}\tilde{\mathbf{C}}_2 + \alpha_{i,3}\tilde{\mathbf{C}}_3, \text{ where } \alpha_{i,1} = 1 - \alpha_{i,2} - \alpha_{i,3}$$

$$\tilde{\mathbf{X}}_i = \tilde{\mathbf{C}}_1 + \alpha_{i,2}(\tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1) + \alpha_{i,3}(\tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1)$$

solve for $\alpha_{i,2}$ and $\alpha_{i,3}$

$$(\tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1)\alpha_{i,2} + (\tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1)\alpha_{i,3} = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\begin{bmatrix} \tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1 \end{bmatrix} \begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \end{bmatrix} = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\mathtt{A} \begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \end{bmatrix} = \mathbf{b}_i, \text{ where } \mathtt{A} = \begin{bmatrix} \tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1 \end{bmatrix}$$

$$\text{and } \mathbf{b}_i = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \end{bmatrix} = \mathtt{A}^\top \mathbf{b}_i, \text{ then } \alpha_{i,1} = 1 - \alpha_{i,2} - \alpha_{i,3}$$

Parameterize 3D points in **world** coordinate frame using control points in **world** coordinate frame

Linear estimation of normalized camera projection matrix

- Control points in camera coordinate frame
  - From parameterization of 3D points, coplanar

Determine the control points

$$\tilde{\mathbf{C}}_{\text{cam},1} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},1} \\ \tilde{Y}^{\text{control}}_{\text{cam},1} \\ \tilde{Z}^{\text{control}}_{\text{cam},1} \end{bmatrix}, \tilde{\mathbf{C}}_{\text{cam},2} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},2} \\ \tilde{Y}^{\text{control}}_{\text{cam},2} \\ \tilde{Z}^{\text{control}}_{\text{cam},2} \end{bmatrix}, \text{ and } \tilde{\mathbf{C}}_{\text{cam},3} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},3} \\ \tilde{Y}^{\text{control}}_{\text{cam},3} \\ \tilde{Z}^{\text{control}}_{\text{cam},3} \end{bmatrix}$$

used to parameterize the i-th point in the camera coordinate frame

$$\tilde{\mathbf{X}}_{\text{cam},i} = \alpha_{i,1}\tilde{\mathbf{C}}_{\text{cam},1} + \alpha_{i,2}\tilde{\mathbf{C}}_{\text{cam},2} + \alpha_{i,3}\tilde{\mathbf{C}}_{\text{cam},3}$$

$$\tilde{\mathbf{X}}_{\text{cam},i} = \sum_{j=1}^{3} \alpha_{i,j}\tilde{\mathbf{C}}_{\text{cam},j}$$

$$\begin{bmatrix} \tilde{X}_{\text{cam},i} \\ \tilde{Y}_{\text{cam},i} \\ \tilde{Z}_{\text{cam},i} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{3} \alpha_{i,j}\tilde{X}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{3} \alpha_{i,j}\tilde{Y}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{3} \alpha_{i,j}\tilde{Z}^{\text{control}}_{\text{cam},j} \end{bmatrix}$$

Linear estimation of normalized camera projection matrix

3D points, coplanar

Note

$$\hat{\tilde{x}}_i = \frac{\hat{x}_i}{\hat{w}_i} \quad \text{and } \hat{\tilde{y}}_i = \frac{\hat{y}_i}{\hat{w}_i}$$
$$\hat{x}_i = \hat{w}_i \hat{\tilde{x}}_i \text{ and } \hat{y}_i = \hat{w}_i \hat{\tilde{y}}_i$$

and recall

$$\hat{\mathbf{x}}_i = \tilde{\mathbf{X}}_{\text{cam},i}$$
$$\begin{bmatrix} \hat{x}_i \\ \hat{y}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \tilde{X}_{\text{cam},i} \\ \tilde{Y}_{\text{cam},i} \\ \tilde{Z}_{\text{cam},i} \end{bmatrix}$$
$$\begin{bmatrix} \hat{w}_i \tilde{\hat{x}}_i \\ \hat{w}_i \tilde{\hat{y}}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{3} \alpha_{i,j} \tilde{X}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{3} \alpha_{i,j} \tilde{Y}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{3} \alpha_{i,j} \tilde{Z}^{\text{control}}_{\text{cam},j} \end{bmatrix}$$

Linear estimation of normalized camera projection matrix

3D points, coplanar

$$\begin{bmatrix} \hat{w}_i \hat{\tilde{x}}_i \\ \hat{w}_i \hat{\tilde{y}}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{3} \alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}} \\ \sum_{j=1}^{3} \alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}} \\ \sum_{j=1}^{3} \alpha_{i,j} \tilde{Z}_{\text{cam},j}^{\text{control}} \end{bmatrix}$$

Substitute the third row into the first row

$$\sum_{j=1}^{3} \alpha_{i,j} \hat{\tilde{x}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} = \sum_{j=1}^{3} \alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}}$$

$$\sum_{j=1}^{3} \left(\alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}} - \alpha_{i,j} \hat{\tilde{x}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} \right) = 0$$

and substitute the third row into the second row

$$\sum_{j=1}^{3} \alpha_{i,j} \hat{\tilde{y}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} = \sum_{j=1}^{3} \alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}}$$

$$\sum_{j=1}^{3} \left(\alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}} - \alpha_{i,j} \hat{\tilde{y}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} \right) = 0$$

Linear estimation of normalized camera projection matrix

3D points, coplanar

Expand both results and put in matrix form

$$\begin{bmatrix} \alpha_{i,1} & 0 & -\alpha_{i,1}\hat{\tilde{x}}_i & \alpha_{i,2} & 0 & -\alpha_{i,2}\hat{\tilde{x}}_i & \alpha_{i,3} & 0 & -\alpha_{i,3}\hat{\tilde{x}}_i \\ 0 & \alpha_{i,1} & -\alpha_{i,1}\hat{\tilde{y}}_i & 0 & \alpha_{i,2} & -\alpha_{i,2}\hat{\tilde{y}}_i & 0 & \alpha_{i,3} & -\alpha_{i,3}\hat{\tilde{y}}_i \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \end{bmatrix} = \mathbf{0} \,\forall\, i$$

$$\begin{bmatrix} \alpha_{1,1} & 0 & -\alpha_{1,1}\hat{\tilde{x}}_1 & \alpha_{1,2} & 0 & -\alpha_{1,2}\hat{\tilde{x}}_1 & \alpha_{1,3} & 0 & -\alpha_{1,3}\hat{\tilde{x}}_1 \\ 0 & \alpha_{1,1} & -\alpha_{1,1}\hat{\tilde{y}}_1 & 0 & \alpha_{1,2} & -\alpha_{1,2}\hat{\tilde{y}}_1 & 0 & \alpha_{1,3} & -\alpha_{1,3}\hat{\tilde{y}}_1 \\ \alpha_{2,1} & 0 & -\alpha_{2,1}\hat{\tilde{x}}_2 & \alpha_{2,2} & 0 & -\alpha_{2,2}\hat{\tilde{x}}_2 & \alpha_{2,3} & 0 & -\alpha_{2,3}\hat{\tilde{x}}_2 \\ 0 & \alpha_{2,1} & -\alpha_{2,1}\hat{\tilde{y}}_2 & 0 & \alpha_{2,2} & -\alpha_{2,2}\hat{\tilde{y}}_2 & 0 & \alpha_{2,3} & -\alpha_{2,3}\hat{\tilde{y}}_2 \\ & & & & & & \vdots & & \\ \alpha_{n,1} & 0 & -\alpha_{n,1}\hat{\tilde{x}}_n & \alpha_{n,2} & 0 & -\alpha_{n,2}\hat{\tilde{x}}_n & \alpha_{n,3} & 0 & -\alpha_{n,3}\hat{\tilde{x}}_n \\ 0 & \alpha_{n,1} & -\alpha_{n,1}\hat{\tilde{y}}_n & 0 & \alpha_{n,2} & -\alpha_{n,2}\hat{\tilde{y}}_n & 0 & \alpha_{n,3} & -\alpha_{n,3}\hat{\tilde{y}}_n \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \end{bmatrix} = \mathbf{0}$$

$$\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}} = \mathbf{0}$$

Linear estimation of normalized camera projection matrix

3D points, coplanar

$$\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}} = \mathbf{0}$$

where

$$\mathtt{M} = \begin{bmatrix} \alpha_{1,1} & 0 & -\alpha_{1,1}\hat{\tilde{x}}_1 & \alpha_{1,2} & 0 & -\alpha_{1,2}\hat{\tilde{x}}_1 & \alpha_{1,3} & 0 & -\alpha_{1,3}\hat{\tilde{x}}_1 \\ 0 & \alpha_{1,1} & -\alpha_{1,1}\hat{\tilde{y}}_1 & 0 & \alpha_{1,2} & -\alpha_{1,2}\hat{\tilde{y}}_1 & 0 & \alpha_{1,3} & -\alpha_{1,3}\hat{\tilde{y}}_1 \\ \alpha_{2,1} & 0 & -\alpha_{2,1}\hat{\tilde{x}}_2 & \alpha_{2,2} & 0 & -\alpha_{2,2}\hat{\tilde{x}}_2 & \alpha_{2,3} & 0 & -\alpha_{2,3}\hat{\tilde{x}}_2 \\ 0 & \alpha_{2,1} & -\alpha_{2,1}\hat{\tilde{y}}_2 & 0 & \alpha_{2,2} & -\alpha_{2,2}\hat{\tilde{y}}_2 & 0 & \alpha_{2,3} & -\alpha_{2,3}\hat{\tilde{y}}_2 \\ & & & & & & \vdots & & \\ \alpha_{n,1} & 0 & -\alpha_{n,1}\hat{\tilde{x}}_i & \alpha_{n,2} & 0 & -\alpha_{n,2}\hat{\tilde{x}}_i & \alpha_{n,3} & 0 & -\alpha_{n,3}\hat{\tilde{x}}_i \\ 0 & \alpha_{n,1} & -\alpha_{n,1}\hat{\tilde{y}}_i & 0 & \alpha_{n,2} & -\alpha_{n,2}\hat{\tilde{y}}_i & 0 & \alpha_{n,3} & -\alpha_{n,3}\hat{\tilde{y}}_i \end{bmatrix}$$

and

$$\mathbf{c}_{\text{cam}}^{\text{control}} = \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \end{bmatrix}$$

solve for $\mathbf{c}_{\text{cam}}^{\text{control}}$

Linear estimation of normalized camera projection matrix

3D points, coplanar

In general, due to error in measurements, the (right) null space of the matrix $\mathtt{M}$ will be empty (i.e., $\mathtt{M}$ will be full rank). Instead, solve for the vector $\mathbf{c}_{\text{cam}}^{\text{control}}$ such that $\|\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}}\|$ is minimized.

$$\mathtt{M} = \mathtt{U}\Sigma\mathtt{V}^{\top}$$

The solution vector $\mathbf{c}_{\text{cam}}^{\text{control}}$ is the column of $\mathtt{V}$ corresponding to the smallest singular value (i.e., the last column of $\mathtt{V}$ (or last row of $\mathtt{V}^{\top}$)).
Note the control points in the camera coordinate frame have been determined up to nonzero scale. As such, the deparameterized, in the camera coordinate frame, parameterizations of 3D points

$$\tilde{\mathbf{X}}_{\text{cam},i} = \alpha_{i,1}\tilde{\mathbf{C}}_{\text{cam},1} + \alpha_{i,2}\tilde{\mathbf{C}}_{\text{cam},2} + \alpha_{i,3}\tilde{\mathbf{C}}_{\text{cam},3}, \text{ where } \alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} = 1 \,\forall\, i$$

are determined up to nonzero scale



Deparameterize to 3D points in **camera** coordinate frame using control points in **camera** coordinate frame

Parameterization of 3D points, general position

Given control points $\tilde{\mathbf{C}}_1, \tilde{\mathbf{C}}_2, \tilde{\mathbf{C}}_3,$ and $\tilde{\mathbf{C}}_4$

$$\tilde{\mathbf{X}}_i = \sum_{j=1}^{4} \alpha_{i,j} \tilde{\mathbf{C}}_j \; \forall \, i$$

$$\tilde{\mathbf{X}}_i = \alpha_{i,1}\tilde{\mathbf{C}}_1 + \alpha_{i,2}\tilde{\mathbf{C}}_2 + \alpha_{i,3}\tilde{\mathbf{C}}_3 + \alpha_{i,4}\tilde{\mathbf{C}}_4, \text{ where } \alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} + \alpha_{i,4} = 1 \, \forall \, i$$

solve for parameterization $(\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}, \alpha_{i,4})$ of i-th point $\tilde{\mathbf{X}}_i$

$$\tilde{\mathbf{X}}_i = (1 - \alpha_{i,2} - \alpha_{i,3} - \alpha_{i,4})\tilde{\mathbf{C}}_1 + \alpha_{i,2}\tilde{\mathbf{C}}_2 + \alpha_{i,3}\tilde{\mathbf{C}}_3 + \alpha_{i,4}\tilde{\mathbf{C}}_4, \text{ where } \alpha_{i,1} = 1 - \alpha_{i,2} - \alpha_{i,3} - \alpha_{i,4}$$

$$\tilde{\mathbf{X}}_i = \tilde{\mathbf{C}}_1 + \alpha_{i,2}(\tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1) + \alpha_{i,3}(\tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1) + \alpha_{i,4}(\tilde{\mathbf{C}}_4 - \tilde{\mathbf{C}}_1)$$

solve for $\alpha_{i,2}$, $\alpha_{i,3}$, and $\alpha_{i,4}$

$$(\tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1)\alpha_{i,2} + (\tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1)\alpha_{i,3} + (\tilde{\mathbf{C}}_4 - \tilde{\mathbf{C}}_1)\alpha_{i,4} = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\begin{bmatrix} \tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_4 - \tilde{\mathbf{C}}_1 \end{bmatrix} \begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \\ \alpha_{i,4} \end{bmatrix} = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\mathtt{A} \begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \\ \alpha_{i,4} \end{bmatrix} = \mathbf{b}_i, \text{ where } \mathtt{A} = \begin{bmatrix} \tilde{\mathbf{C}}_2 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_3 - \tilde{\mathbf{C}}_1 & \tilde{\mathbf{C}}_4 - \tilde{\mathbf{C}}_1 \end{bmatrix}$$

$$\text{and } \mathbf{b}_i = \tilde{\mathbf{X}}_i - \tilde{\mathbf{C}}_1$$

$$\begin{bmatrix} \alpha_{i,2} \\ \alpha_{i,3} \\ \alpha_{i,4} \end{bmatrix} = \mathtt{A}^\top \mathbf{b}_i, \text{ then } \alpha_{i,1} = 1 - \alpha_{i,2} - \alpha_{i,3} - \alpha_{i,4}$$

Parameterize 3D points in **world** coordinate frame using control points in **world** coordinate frame

Linear estimation of normalized camera projection matrix

- Control points in camera coordinate frame
  - From parameterization of 3D points, general position

Determine the control points

$$\tilde{\mathbf{C}}_{\text{cam},1} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},1} \\ \tilde{Y}^{\text{control}}_{\text{cam},1} \\ \tilde{Z}^{\text{control}}_{\text{cam},1} \end{bmatrix}, \tilde{\mathbf{C}}_{\text{cam},2} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},2} \\ \tilde{Y}^{\text{control}}_{\text{cam},2} \\ \tilde{Z}^{\text{control}}_{\text{cam},2} \end{bmatrix}, \tilde{\mathbf{C}}_{\text{cam},3} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},3} \\ \tilde{Y}^{\text{control}}_{\text{cam},3} \\ \tilde{Z}^{\text{control}}_{\text{cam},3} \end{bmatrix}, \text{ and } \tilde{\mathbf{C}}_{\text{cam},4} = \begin{bmatrix} \tilde{X}^{\text{control}}_{\text{cam},4} \\ \tilde{Y}^{\text{control}}_{\text{cam},4} \\ \tilde{Z}^{\text{control}}_{\text{cam},4} \end{bmatrix}$$

used to parameterize the i-th point in camera coordinate frame

$$\tilde{\mathbf{X}}_{\text{cam},i} = \alpha_{i,1}\tilde{\mathbf{C}}_{\text{cam},1} + \alpha_{i,2}\tilde{\mathbf{C}}_{\text{cam},2} + \alpha_{i,3}\tilde{\mathbf{C}}_{\text{cam},3} + \alpha_{i,4}\tilde{\mathbf{C}}_{\text{cam},4}$$

$$\tilde{\mathbf{X}}_{\text{cam},i} = \sum_{j=1}^{4} \alpha_{i,j}\tilde{\mathbf{C}}_{\text{cam},j}$$

$$\begin{bmatrix} \tilde{X}_{\text{cam},i} \\ \tilde{Y}_{\text{cam},i} \\ \tilde{Z}_{\text{cam},i} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{4} \alpha_{i,j}\tilde{X}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{4} \alpha_{i,j}\tilde{Y}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{4} \alpha_{i,j}\tilde{Z}^{\text{control}}_{\text{cam},j} \end{bmatrix}$$

Linear estimation of normalized camera projection matrix

3D points, general position

Note

$$\hat{\tilde{x}}_i = \frac{\hat{x}_i}{\hat{w}_i} \quad \text{and } \hat{\tilde{y}}_i = \frac{\hat{y}_i}{\hat{w}_i}$$
$$\hat{x}_i = \hat{w}_i \hat{\tilde{x}}_i \text{ and } \hat{y}_i = \hat{w}_i \hat{\tilde{y}}_i$$

and recall

$$\hat{\mathbf{x}}_i = \tilde{\mathbf{X}}_{\text{cam},i}$$
$$\begin{bmatrix} \hat{x}_i \\ \hat{y}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \tilde{X}_{\text{cam},i} \\ \tilde{Y}_{\text{cam},i} \\ \tilde{Z}_{\text{cam},i} \end{bmatrix}$$
$$\begin{bmatrix} \hat{w}_i \hat{\tilde{x}}_i \\ \hat{w}_i \hat{\tilde{y}}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{4} \alpha_{i,j} \tilde{X}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{4} \alpha_{i,j} \tilde{Y}^{\text{control}}_{\text{cam},j} \\ \sum_{j=1}^{4} \alpha_{i,j} \tilde{Z}^{\text{control}}_{\text{cam},j} \end{bmatrix}$$

Linear estimation of normalized camera projection matrix

3D points, general position

$$\begin{bmatrix} \hat{w}_i \hat{\tilde{x}}_i \\ \hat{w}_i \hat{\tilde{y}}_i \\ \hat{w}_i \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{4} \alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}} \\ \sum_{j=1}^{4} \alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}} \\ \sum_{j=1}^{4} \alpha_{i,j} \tilde{Z}_{\text{cam},j}^{\text{control}} \end{bmatrix}$$

Substitute the third row into the first row

$$\sum_{j=1}^{4} \alpha_{i,j} \hat{\tilde{x}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} = \sum_{j=1}^{4} \alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}}$$

$$\sum_{j=1}^{4} \left(\alpha_{i,j} \tilde{X}_{\text{cam},j}^{\text{control}} - \alpha_{i,j} \hat{\tilde{x}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} \right) = 0$$

and substitute the third row into the second row

$$\sum_{j=1}^{4} \alpha_{i,j} \hat{\tilde{y}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} = \sum_{j=1}^{4} \alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}}$$

$$\sum_{j=1}^{4} \left(\alpha_{i,j} \tilde{Y}_{\text{cam},j}^{\text{control}} - \alpha_{i,j} \hat{\tilde{y}}_i \tilde{Z}_{\text{cam},j}^{\text{control}} \right) = 0$$

Linear estimation of normalized camera projection matrix

3D points, general position

Expand both results and put in matrix form

$$\begin{bmatrix} \alpha_{i,1} & 0 & -\alpha_{i,1}\hat{\tilde{x}}_i & \alpha_{i,2} & 0 & -\alpha_{i,2}\hat{\tilde{x}}_i & \alpha_{i,3} & 0 & -\alpha_{i,3}\hat{\tilde{x}}_i & \alpha_{i,4} & 0 & -\alpha_{i,4}\hat{\tilde{x}}_i \\ 0 & \alpha_{i,1} & -\alpha_{i,1}\hat{\tilde{y}}_i & 0 & \alpha_{i,2} & -\alpha_{i,2}\hat{\tilde{y}}_i & 0 & \alpha_{i,3} & -\alpha_{i,3}\hat{\tilde{y}}_i & 0 & \alpha_{i,4} & -\alpha_{i,4}\hat{\tilde{y}}_i \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \\ \tilde{\mathbf{C}}_{\text{cam},4} \end{bmatrix} = \mathbf{0} \,\forall\, i$$

$$\begin{bmatrix} \alpha_{1,1} & 0 & -\alpha_{1,1}\hat{\tilde{x}}_1 & \alpha_{1,2} & 0 & -\alpha_{1,2}\hat{\tilde{x}}_1 & \alpha_{1,3} & 0 & -\alpha_{1,3}\hat{\tilde{x}}_1 & \alpha_{1,4} & 0 & -\alpha_{1,4}\hat{\tilde{x}}_1 \\ 0 & \alpha_{1,1} & -\alpha_{1,1}\hat{\tilde{y}}_1 & 0 & \alpha_{1,2} & -\alpha_{1,2}\hat{\tilde{y}}_1 & 0 & \alpha_{1,3} & -\alpha_{1,3}\hat{\tilde{y}}_1 & 0 & \alpha_{1,4} & -\alpha_{1,4}\hat{\tilde{y}}_1 \\ \alpha_{2,1} & 0 & -\alpha_{2,1}\hat{\tilde{x}}_2 & \alpha_{2,2} & 0 & -\alpha_{2,2}\hat{\tilde{x}}_2 & \alpha_{2,3} & 0 & -\alpha_{2,3}\hat{\tilde{x}}_2 & \alpha_{2,4} & 0 & -\alpha_{2,4}\hat{\tilde{x}}_2 \\ 0 & \alpha_{2,1} & -\alpha_{2,1}\hat{\tilde{y}}_2 & 0 & \alpha_{2,2} & -\alpha_{2,2}\hat{\tilde{y}}_2 & 0 & \alpha_{2,3} & -\alpha_{2,3}\hat{\tilde{y}}_2 & 0 & \alpha_{2,4} & -\alpha_{2,4}\hat{\tilde{y}}_2 \\ & & & & & & \vdots & & & & & \\ \alpha_{n,1} & 0 & -\alpha_{n,1}\hat{\tilde{x}}_n & \alpha_{n,2} & 0 & -\alpha_{n,2}\hat{\tilde{x}}_n & \alpha_{n,3} & 0 & -\alpha_{n,3}\hat{\tilde{x}}_n & \alpha_{n,4} & 0 & -\alpha_{n,4}\hat{\tilde{x}}_n \\ 0 & \alpha_{n,1} & -\alpha_{n,1}\hat{\tilde{y}}_n & 0 & \alpha_{n,2} & -\alpha_{n,2}\hat{\tilde{y}}_n & 0 & \alpha_{n,3} & -\alpha_{n,3}\hat{\tilde{y}}_n & 0 & \alpha_{n,4} & -\alpha_{n,4}\hat{\tilde{y}}_n \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \\ \tilde{\mathbf{C}}_{\text{cam},4} \end{bmatrix} = \mathbf{0}$$

$$\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}} = \mathbf{0}$$

Linear estimation of normalized camera projection matrix

3D points, general position

$$\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}} = \mathbf{0}$$

where

$$\mathtt{M} = \begin{bmatrix}
\alpha_{1,1} & 0 & -\alpha_{1,1}\hat{\tilde{x}}_1 & \alpha_{1,2} & 0 & -\alpha_{1,2}\hat{\tilde{x}}_1 & \alpha_{1,3} & 0 & -\alpha_{1,3}\hat{\tilde{x}}_1 & \alpha_{1,4} & 0 & -\alpha_{1,4}\hat{\tilde{x}}_1 \\
0 & \alpha_{1,1} & -\alpha_{1,1}\hat{\tilde{y}}_1 & 0 & \alpha_{1,2} & -\alpha_{1,2}\hat{\tilde{y}}_1 & 0 & \alpha_{1,3} & -\alpha_{1,3}\hat{\tilde{y}}_1 & 0 & \alpha_{1,4} & -\alpha_{1,4}\hat{\tilde{y}}_1 \\
\alpha_{2,1} & 0 & -\alpha_{2,1}\hat{\tilde{x}}_2 & \alpha_{2,2} & 0 & -\alpha_{2,2}\hat{\tilde{x}}_2 & \alpha_{2,3} & 0 & -\alpha_{2,3}\hat{\tilde{x}}_2 & \alpha_{2,4} & 0 & -\alpha_{2,4}\hat{\tilde{x}}_2 \\
0 & \alpha_{2,1} & -\alpha_{2,1}\hat{\tilde{y}}_2 & 0 & \alpha_{2,2} & -\alpha_{2,2}\hat{\tilde{y}}_2 & 0 & \alpha_{2,3} & -\alpha_{2,3}\hat{\tilde{y}}_2 & 0 & \alpha_{2,4} & -\alpha_{2,4}\hat{\tilde{y}}_2 \\
 & & & & & & \vdots & & & & & \\
\alpha_{n,1} & 0 & -\alpha_{n,1}\hat{\tilde{x}}_i & \alpha_{n,2} & 0 & -\alpha_{n,2}\hat{\tilde{x}}_i & \alpha_{n,3} & 0 & -\alpha_{n,3}\hat{\tilde{x}}_i & \alpha_{n,4} & 0 & -\alpha_{n,4}\hat{\tilde{x}}_i \\
0 & \alpha_{n,1} & -\alpha_{n,1}\hat{\tilde{y}}_i & 0 & \alpha_{n,2} & -\alpha_{n,2}\hat{\tilde{y}}_i & 0 & \alpha_{n,3} & -\alpha_{n,3}\hat{\tilde{y}}_i & 0 & \alpha_{n,4} & -\alpha_{n,4}\hat{\tilde{y}}_i
\end{bmatrix}$$

and

$$\mathbf{c}_{\text{cam}}^{\text{control}} = \begin{bmatrix} \tilde{\mathbf{C}}_{\text{cam},1} \\ \tilde{\mathbf{C}}_{\text{cam},2} \\ \tilde{\mathbf{C}}_{\text{cam},3} \\ \tilde{\mathbf{C}}_{\text{cam},4} \end{bmatrix}$$

solve for $\mathbf{c}_{\text{cam}}^{\text{control}}$

Linear estimation of normalized camera projection matrix

3D points, general position

In general, due to error in measurements, the (right) null space of the matrix $\mathtt{M}$ will be empty (i.e., $\mathtt{M}$ will be full rank). Instead, solve for the vector $\mathbf{c}_{\text{cam}}^{\text{control}}$ such that $\|\mathtt{M}\mathbf{c}_{\text{cam}}^{\text{control}}\|$ is minimized.

$$\mathtt{M} = \mathtt{U}\Sigma\mathtt{V}^{\top}$$

The solution vector $\mathbf{c}_{\text{cam}}^{\text{control}}$ is the column of $\mathtt{V}$ corresponding to the smallest singular value (i.e., the last column of $\mathtt{V}$ (or last row of $\mathtt{V}^{\top}$)).
Note the control points in the camera coordinate frame have been determined up to nonzero scale. As such, the deparamterized, in the camera coordinate frame, parameterizations of 3D points

$$\tilde{\mathbf{X}}_{\text{cam},i} = \alpha_{i,1}\tilde{\mathbf{C}}_{\text{cam},1} + \alpha_{i,2}\tilde{\mathbf{C}}_{\text{cam},2} + \alpha_{i,3}\tilde{\mathbf{C}}_{\text{cam},3} + \alpha_{i,4}\tilde{\mathbf{C}}_{\text{cam},4}, \text{ where } \alpha_{i,1} + \alpha_{i,2} + \alpha_{i,3} + \alpha_{i,4} = 1 \,\forall\, i$$

are determined up to nonzero scale

Deparameterize to 3D points in **camera** coordinate frame using control points in **camera** coordinate frame

Linear estimation of normalized camera projection matrix

- Scale 3D points in camera coordinate frame

1. Calculate the total variance $\sigma^2_{\tilde{\mathbf{X}}} = \mathrm{Tr}(\Sigma_{\tilde{\mathbf{X}}})$ of the set of 3D points in the world coordinate frame
2. Calculate the mean $\mu_{\tilde{Z}_{\mathrm{cam}}}$ of the $\tilde{Z}_{\mathrm{cam}}$ coordinate and the total variance $\sigma^2_{\tilde{\mathbf{X}}_{\mathrm{cam}}} = \sigma^2_{\tilde{X}_{\mathrm{cam}}} + \sigma^2_{\tilde{Y}_{\mathrm{cam}}} + \sigma^2_{\tilde{Z}_{\mathrm{cam}}}$ of the set of deparameterized 3D points in the camera coordinate frame
3. Determine scale

$$\beta = \begin{cases} \sqrt{\frac{\sigma^2_{\tilde{\mathbf{X}}}}{\sigma^2_{\tilde{\mathbf{X}}_{\mathrm{cam}}}}} & \text{if } \mu_{\tilde{Z}_{\mathrm{cam}}} \geq 0 \\ -\sqrt{\frac{\sigma^2_{\tilde{\mathbf{X}}}}{\sigma^2_{\tilde{\mathbf{X}}_{\mathrm{cam}}}}} & \text{otherwise} \end{cases}$$

   such that $\tilde{Z}_{\mathrm{cam}}$ coordinates are positive (i.e., in front of the camera)

4. Scale 3D points in the camera coordinate frame

$$\tilde{\mathbf{X}}_{\mathrm{cam},i} = \beta \tilde{\mathbf{X}}_{\mathrm{cam},i} \, \forall \, i$$

Estimate 3D Euclidean transformation from points in the world coordinate frame to points in the camera coordinate frame

$\tilde{\mathbf{X}}_{\text{cam},i} = \mathtt{R}\tilde{\mathbf{X}}_i + \mathbf{t} \,\forall\, i$, solve for $\mathtt{R}$ and $\mathbf{t}$ using $\tilde{\mathbf{X}}_{\text{cam},i} \leftrightarrow \tilde{\mathbf{X}}_i$

1. Calculate the mean vectors $\boldsymbol{\mu}_{\tilde{\mathbf{X}}}$ and $\boldsymbol{\mu}_{\tilde{\mathbf{X}}_{\text{cam}}}$ (i.e., centroids) of the set of 3D points in world coordinate frame and camera coordinate frame, respectively

2. Compute the matrix

$$\mathtt{S} = \sum_i \left(\tilde{\mathbf{X}}_{\text{cam},i} - \boldsymbol{\mu}_{\tilde{\mathbf{X}}_{\text{cam}}}\right)\left(\tilde{\mathbf{X}}_i - \boldsymbol{\mu}_{\tilde{\mathbf{X}}}\right)^\top$$

   and its singular value decomposition $\mathtt{S} = \mathtt{U}\Sigma\mathtt{V}^\top$

3. The camera rotation matrix

$$\mathtt{R} = \begin{cases} \mathtt{U}\,\mathrm{diag}(1,1,-1)\mathtt{V}^\top & \text{if } \det(\mathtt{U})\det(\mathtt{V}) < 0 \\ \mathtt{U}\mathtt{V}^\top & \text{otherwise} \end{cases}$$

   such that $\det(\mathtt{R}) = +1$

4. The camera translation vector $\mathbf{t} = \boldsymbol{\mu}_{\tilde{\mathbf{X}}_{\text{cam}}} - \mathtt{R}\boldsymbol{\mu}_{\tilde{\mathbf{X}}}$

Next lecture

- Estimation of camera pose (calibrated camera), nonlinear