

# 3D Computer Vision

Radim Šára    Martin Matoušek

Center for Machine Perception  
Department of Cybernetics  
Faculty of Electrical Engineering  
Czech Technical University in Prague

<https://cw.fel.cvut.cz/wiki/courses/tdv/start>

<http://cmp.felk.cvut.cz>

<mailto:sara@cmp.felk.cvut.cz>

phone ext. 7203

rev. October 11, 2022

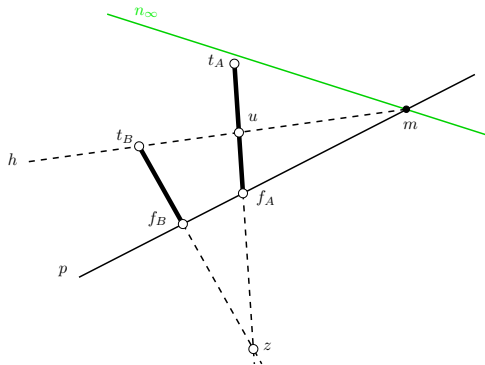
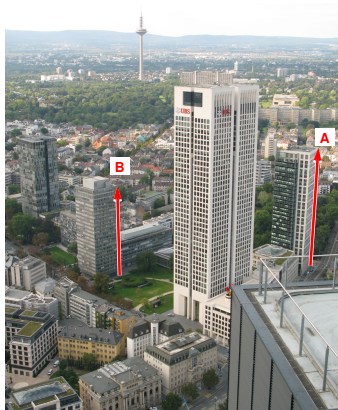


Open Informatics Master's Course

# Homework Problem

⊗ H2; 3pt: What is the ratio of heights of Building A to Building B?

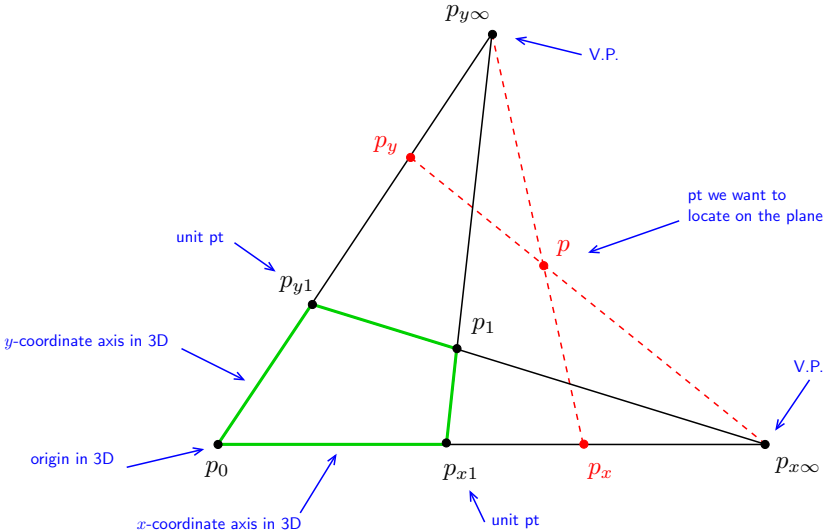
- expected: conceptual solution; use notation from this figure
- deadline: LD+2 weeks



## Hints

1. What are the interesting properties of line  $h$  connecting the top  $t_B$  of Building B with the point  $m$  at which the horizon intersects the line  $p$  joining the feet  $f_A, f_B$  of both buildings? [1 point]
2. How do we actually get the horizon  $n_\infty$ ? (we do not see it directly, there are some hills there...) [1 point]
3. Give a formula for measuring the length ratio. Make sure you distinguish points in 3D from their images. [formula = 1 point]

# 2D Projective Coordinates



$$[P_x] = [P_0 \ P_{x1} \ P_x \ P_{x\infty}]$$

$$[P_y] = [P_0 \ P_{y1} \ P_y \ P_{y\infty}]$$

## Application: Measuring on the Floor (Wall, etc)



San Giovanni in Laterano, Rome

- measuring distances on the floor in terms of tile units
- what are the dimensions of the seal? Is it circular (assuming square tiles)?
- needs no explicit camera calibration

because we can see the calibrating object (vanishing points)

## Computing with a Single Camera

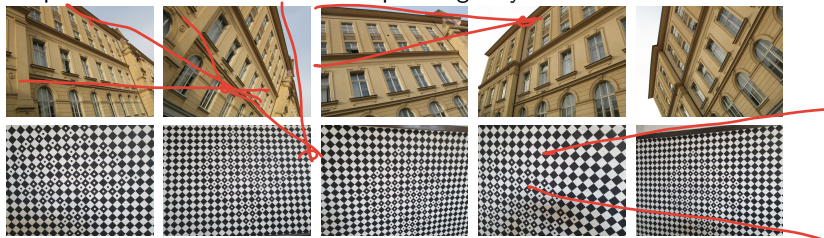
- 3.1 Calibration: Internal Camera Parameters from Vanishing Points and Lines
- 3.2 Camera Resection: Projection Matrix from 6 Known Points
- 3.3 Exterior Orientation: Camera Rotation and Translation from 3 Known Points
- 3.4 Relative Orientation Problem: Rotation and Translation between Two Point Sets

### covered by

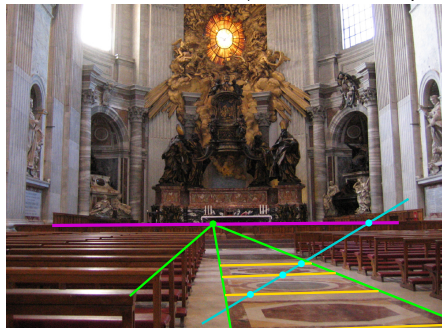
- [1] [H&Z] Secs: 8.6, 7.1, 22.1
- [2] Fischler, M.A. and Bolles, R.C. . Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24(6):381–395, 1981
- [3] [Golub & van Loan 2013, Sec. 2.5]

# Obtaining Vanishing Points and Lines

- orthogonal direction pairs can be collected from multiple images by camera rotation



- vanishing line can be obtained from vanishing points and/or regularities (→49)



## ► Camera Calibration from Vanishing Points and Lines

**Problem:** Given finite vanishing points and/or vanishing lines, compute  $\mathbf{K}$

$$\mathbf{Q} = \mathbf{K}\mathbf{R}$$

$$\lambda_i, \mu_{ij} \neq 0 \neq \infty \quad (2)$$

$$\begin{aligned} \Downarrow \quad \mathbf{d}_i &= \lambda_i \mathbf{Q}^{-1} \mathbf{v}_i, & i = 1, 2, 3 & \rightarrow 43 \\ \Downarrow \quad \mathbf{p}_{ij} &= \mu_{ij} \mathbf{Q}^T \mathbf{n}_{ij}, & i, j = 1, 2, 3, i \neq j & \rightarrow 39 \end{aligned}$$

- method: eliminate  $\lambda_i, \mu_{ij}, \mathbf{R}$  from (2) and solve for  $\mathbf{K}$ .

**Configurations allowing elimination of  $\mathbf{R}$**

1. orthogonal rays  $\mathbf{d}_1 \perp \mathbf{d}_2$  in space then

$$0 = \mathbf{d}_1^T \mathbf{d}_2 = \mathbf{v}_1^T \mathbf{Q}^{-T} \mathbf{Q}^{-1} \mathbf{v}_2 = \mathbf{v}_1^T \underbrace{(\mathbf{K}\mathbf{K}^T)^{-1}}_{\omega \text{ (IAC)}} \mathbf{v}_2$$

2. orthogonal planes  $\mathbf{p}_{ij} \perp \mathbf{p}_{ik}$  in space

$$0 = \mathbf{p}_{ij}^T \mathbf{p}_{ik} = \mathbf{n}_{ij}^T \mathbf{Q}\mathbf{Q}^T \mathbf{n}_{ik} = \mathbf{n}_{ij}^T \omega^{-1} \mathbf{n}_{ik}$$

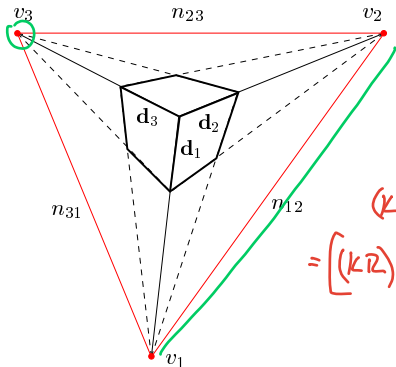
3. orthogonal ray and plane  $\mathbf{d}_k \parallel \mathbf{p}_{ij}, k \neq i, j$

normal parallel to optical ray

$$\mathbf{p}_{ij} \simeq \mathbf{d}_k \Rightarrow \mathbf{Q}^T \mathbf{n}_{ij} = \frac{\lambda_i}{\mu_{ij}} \mathbf{Q}^{-1} \mathbf{v}_k \Rightarrow \mathbf{n}_{ij} = \varkappa \mathbf{Q}^{-T} \mathbf{Q}^{-1} \mathbf{v}_k = \varkappa \omega \mathbf{v}_k, \quad \varkappa \neq 0$$

- $n_{ij}$  may be constructed from non-orthogonal  $v_i$  and  $v_j$ , e.g. using the cross-ratio
- $\omega$  is a symmetric, positive definite  $3 \times 3$  matrix
- equations are quadratic in  $\mathbf{K}$  but linear in  $\omega$

IAC = Image of Absolute Conic



$$= \begin{bmatrix} (\mathbf{K}\mathbf{R})^T & (\mathbf{K}\mathbf{R})^T \\ (\mathbf{K}\mathbf{R})^T & (\mathbf{K}\mathbf{R})^T \end{bmatrix}$$

configuration	equation	# constraints
(3) orthogonal vanishing points	$\mathbf{v}_i^T \boldsymbol{\omega} \mathbf{v}_j = 0$	1
(4) orthogonal vanishing lines	<del><math>\mathbf{n}_{ij}^T \boldsymbol{\omega}^{-1} \mathbf{n}_{ik} = 0</math></del>	1
(5) vanishing points orthogonal to vanishing lines	$\mathbf{n}_{ij} = \times \boldsymbol{\omega} \mathbf{v}_k$	2
(6) orthogonal image raster $\theta = \pi/2$	$\omega_{12} = \omega_{21} = 0$	1
(7) unit aspect $a = 1$ when $\theta = \pi/2$	$\omega_{11} - \omega_{22} = 0$	1
(8) known principal point $u_0 = v_0 = 0$	$\omega_{13} = \omega_{31} = \omega_{23} = \omega_{32} = 0$	2

- these are homogeneous linear equations for the 5 parameters in  $\boldsymbol{\omega}$  or  $\boldsymbol{\omega}^{-1}$  in the form  $\mathbf{D}\mathbf{w} = \mathbf{0}$   
x can be eliminated from (5)
- we need at least 5 constraints for full  $\boldsymbol{\omega}$
- we get  $\mathbf{K}$  from  $\boldsymbol{\omega}^{-1} = \mathbf{K}\mathbf{K}^T$  by Choleski decomposition  
chol(inv(omg)) symmetric  $3 \times 3$   
 the decomposition returns a positive definite upper triangular matrix  
 one avoids solving an explicit set of quadratic equations for the parameters in  $\mathbf{K}$



## Examples

Assuming orthogonal raster, unit aspect (ORUA):  $\theta = \pi/2$ ,  $a = 1$

$$\omega \simeq \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ -u_0 & -v_0 & f^2 + u_0^2 + v_0^2 \end{bmatrix}$$

$$\omega \simeq \begin{bmatrix} 1 & & \\ & 1 & \\ & & f^2 \end{bmatrix}$$

### Ex 1:

Assuming ORUA and known  $m_0 = (u_0, v_0)$ , two finite orthogonal vanishing points give  $f$

$$\mathbf{v}_1^\top \omega \mathbf{v}_2 = 0 \quad \Rightarrow \quad f^2 = |(\mathbf{v}_1 - \mathbf{m}_0)^\top (\mathbf{v}_2 - \mathbf{m}_0)|$$

in this formula,  $\mathbf{v}_{1,2}$ ,  $\mathbf{m}_0$  are Cartesian (not homogeneous)!

### Ex 2:

Non-orthogonal vanishing points  $\mathbf{v}_i$ ,  $\mathbf{v}_j$ , known angle  $\phi$ :  $\cos \phi = \frac{\mathbf{v}_i^\top \omega \mathbf{v}_j}{\sqrt{\mathbf{v}_i^\top \omega \mathbf{v}_i} \sqrt{\mathbf{v}_j^\top \omega \mathbf{v}_j}}$

- leads to polynomial equations
- e.g. ORUA and  $u_0 = v_0 = 0$  gives

$$(f^2 + \mathbf{v}_i^\top \mathbf{v}_j)^2 = (f^2 + \|\mathbf{v}_i\|^2) \cdot (f^2 + \|\mathbf{v}_j\|^2) \cdot \cos^2 \phi$$

## ► Camera Orientation from Two Finite Vanishing Points

**Problem:** Given  $\mathbf{K}$  and two vanishing points corresponding to two known orthogonal directions  $\mathbf{d}_1, \mathbf{d}_2$ , compute camera orientation  $\mathbf{R}$  with respect to the plane.

- 3D coordinate system choice, e.g.:

$$\mathbf{d}_1 = (1, 0, 0), \quad \mathbf{d}_2 = (0, 1, 0)$$

- we know that

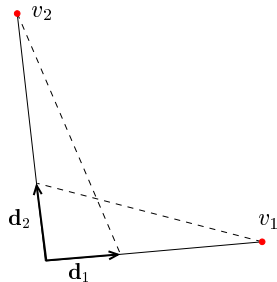
$$\mathbf{d}_i \simeq \mathbf{Q}^{-1} \mathbf{v}_i = (\mathbf{K}\mathbf{R})^{-1} \mathbf{v}_i = \mathbf{R}^{-1} \underbrace{\mathbf{K}^{-1} \mathbf{v}_i}_{\mathbf{w}_i}$$

$\mathbf{R} \mathbf{d}_i \simeq \mathbf{w}_i \cdot \lambda_i$        $\mathbf{v}_i \simeq \mathbf{w}_i = \mathbf{R}^T \mathbf{d}_i$

- knowing  $\mathbf{d}_{1,2}$  we conclude that  $\mathbf{w}_i / \|\mathbf{w}_i\|$  is the  $i$ -th column  $\mathbf{r}_i$  of  $\mathbf{R}$
- the third column is orthogonal:  $\mathbf{r}_3 \simeq \mathbf{r}_1 \times \mathbf{r}_2$

$$\mathbf{R} = \begin{bmatrix} \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} & \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} & \frac{\mathbf{w}_1 \times \mathbf{w}_2}{\|\mathbf{w}_1 \times \mathbf{w}_2\|} \end{bmatrix}$$

- we have to care about the signs  $\pm \mathbf{w}_i$  (such that  $\det \mathbf{R} = 1$ )

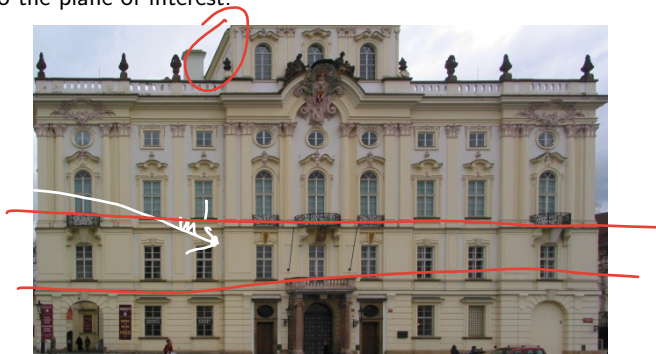
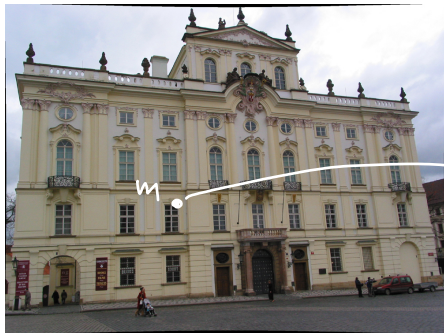


some suitable scenes



# Application: Planar Rectification

**Principle:** Rotate camera (image plane) parallel to the plane of interest.



$$K(KR)^{-1}$$

$$\underline{m} \simeq \underline{KR} [I \quad -C] \underline{X}$$

$$\underline{KR}^{-1} \underline{KR} \underline{m}' \simeq \underline{K}(\underline{KR})^{-1} \underline{m} = \underline{KR}^T \underline{K}^{-1} \underline{m} = \underline{H} \underline{m}$$

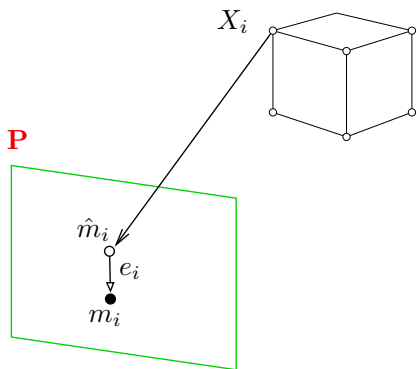
$$\underline{m}' \simeq \underline{K} [I \quad -C] \underline{X}$$

- $\mathbf{H}$  is the rectifying homography
- both  $\mathbf{K}$  and  $\mathbf{R}$  can be calibrated from two finite vanishing points
- not possible when one of them is (or both are) infinite
- without ORUA we would need 4 additional views to calibrate  $\mathbf{K}$  as on →54

assuming ORUA →57

## ► Camera Resection

Camera calibration and orientation from a known set of  $k \geq 6$  reference points and their images  $\{(X_i, m_i)\}_{i=1}^6$ .



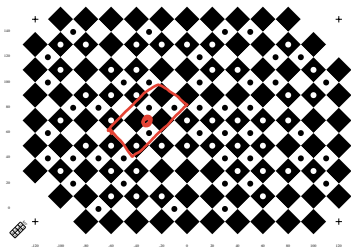
- $X_i$  are considered exact
- $m_i$  is a measurement subject to detection error

$$\mathbf{m}_i = \hat{\mathbf{m}}_i + \mathbf{e}_i \quad \text{Cartesian}$$

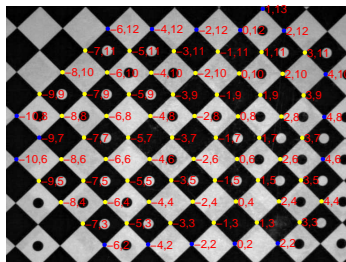
- where  $\lambda_i \hat{\mathbf{m}}_i = \mathbf{P} \mathbf{X}_i$

# Resection Targets

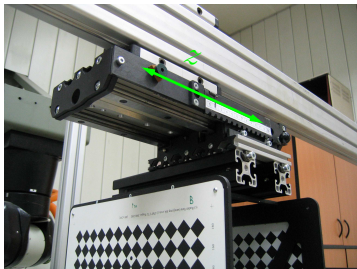
01100110



calibration chart



automatic calibration point detection  
based on a distributed bitcode ( $2 \times 4 = 8$  bits)



resection target with translation stage

- target translated at least once
- by a calibrated (known) translation
- $X_i$  point locations looked up in a table based on their bitcode

## ► The Minimal Problem for Camera Resection

**Problem:** Given  $k = 6$  corresponding pairs  $\{(X_i, m_i)\}_{i=1}^k$ , find  $\mathbf{P}$

$$\lambda_i \underline{\mathbf{m}}_i = \mathbf{P} \underline{\mathbf{X}}_i, \quad \mathbf{P} = \begin{bmatrix} \mathbf{q}_1^\top & q_{14} \\ \mathbf{q}_2^\top & q_{24} \\ \mathbf{q}_3^\top & q_{34} \end{bmatrix}$$

$$\underline{\mathbf{X}}_i = (x_i, y_i, z_i, 1), \quad i = 1, 2, \dots, k, \quad k = 6$$

$$\underline{\mathbf{m}}_i = (u_i, v_i, 1), \quad \lambda_i \in \mathbb{R}, \quad \lambda_i \neq 0, \quad |\lambda_i| < \infty$$

easily modifiable for infinite points  $X_i$  but be aware of  $\rightarrow 64$

expanded:  $\lambda_i u_i = \mathbf{q}_1^\top \mathbf{X}_i + q_{14}, \quad \lambda_i v_i = \mathbf{q}_2^\top \mathbf{X}_i + q_{24}, \quad \lambda_i = \mathbf{q}_3^\top \mathbf{X}_i + q_{34}$

after elimination of  $\lambda_i$ :  $(\mathbf{q}_3^\top \mathbf{X}_i + q_{34})u_i = \mathbf{q}_1^\top \mathbf{X}_i + q_{14}, \quad (\mathbf{q}_3^\top \mathbf{X}_i + q_{34})v_i = \mathbf{q}_2^\top \mathbf{X}_i + q_{24}$

Then

$$\mathbf{A} \mathbf{q} = \begin{bmatrix} \mathbf{X}_1^\top & 1 & \mathbf{0}^\top & 0 & -u_1 \mathbf{X}_1^\top & -u_1 \\ \mathbf{0}^\top & 0 & \mathbf{X}_1^\top & 1 & -v_1 \mathbf{X}_1^\top & -v_1 \\ \vdots & & & & & \\ \mathbf{X}_k^\top & 1 & \mathbf{0}^\top & 0 & -u_k \mathbf{X}_k^\top & -u_k \\ \mathbf{0}^\top & 0 & \mathbf{X}_k^\top & 1 & -v_k \mathbf{X}_k^\top & -v_k \end{bmatrix} \cdot \begin{bmatrix} \mathbf{q}_1 \\ q_{14} \\ \mathbf{q}_2 \\ q_{24} \\ \mathbf{q}_3 \\ q_{34} \end{bmatrix} = \mathbf{0} \quad (9)$$

$= \mathbb{R}^{12}$

$2 \times 12$

- we need 11 independent parameters for  $\mathbf{P}$
- $\mathbf{A} \in \mathbb{R}^{2k, 12}$ ,  $\mathbf{q} \in \mathbb{R}^{12}$
- 6 points in a general position give  $\text{rank } \mathbf{A} = 12$  and there is no (non-trivial) null space
- drop one row to get rank-11 matrix, then the basis vector of the null space of  $\mathbf{A}$  gives  $\mathbf{q}$

## ► The Jack-Knife Solution for $k = 6$

- given the 6 correspondences, we have 12 equations for the 11 parameters
- can we use all the information present in the 6 points?

### Jack-knife estimation

1.  $n := 0$
2. for  $i = 1, 2, \dots, 2k$  do
  - a) delete  $i$ -th row from  $\mathbf{A}$ , this gives  $\mathbf{A}_i$
  - b) if  $\dim \text{null } \mathbf{A}_i > 1$  continue with the next  $i$**
  - c)  $n := n + 1$
  - d) compute the right null-space  $\mathbf{q}_i$  of  $\mathbf{A}_i$
  - e)  $\hat{\mathbf{q}}_i := \mathbf{q}_i$  normalized to  $q_{34} = 1$  and dimension-reduced
3. from all  $n$  vectors  $\hat{\mathbf{q}}_i$  collected in Step 2.e compute

$$\mathbf{q} = \frac{1}{n} \sum_{i=1}^n \hat{\mathbf{q}}_i, \quad \text{var}[\mathbf{q}] = \frac{n-1}{n} \text{diag} \sum_{i=1}^n (\hat{\mathbf{q}}_i - \mathbf{q})(\hat{\mathbf{q}}_i - \mathbf{q})^\top$$

regular for  $n \geq 11$   
variance of the sample mean

- have a solution + an error estimate, per individual elements of  $\mathbf{P}$  (except  $P_{34}$ )
- at least 5 points must be in a general position ( $\rightarrow 64$ )
- large error indicates near degeneracy
- computation not efficient with  $k > 6$  points, needs  $\binom{2k}{11}$  draws, e.g.  $k = 7 \Rightarrow 364$  draws
- better error estimation method: decompose  $\mathbf{P}_i$  to  $\mathbf{K}_i, \mathbf{R}_i, \mathbf{t}_i$  ( $\rightarrow 33$ ), represent  $\mathbf{R}_i$  with 3 parameters (e.g. Euler angles, or in exponential map representation  $\rightarrow 144$ ) and compute the errors for the parameters
- even better: use the SE(3) Lie group for  $(\mathbf{R}_i, \mathbf{t}_i)$  and average its Lie-algebraic representations



e.g. by 'economy-size' SVD  
assuming finite cam. with  $P_{3,4} = 1$

## ► Degenerate (Critical) Configurations for Camera Resection

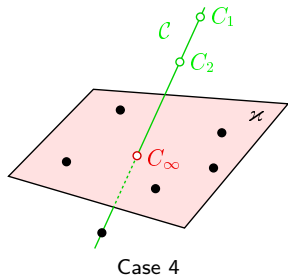
Let  $\mathcal{X} = \{X_i; i = 1, \dots\}$  be a set of points and  $\mathbf{P}_1 \neq \mathbf{P}_j$  be two regular (rank-3) cameras. Then two configurations  $(\mathbf{P}_1, \mathcal{X})$  and  $(\mathbf{P}_j, \mathcal{X})$  are image-equivalent if

$$\mathbf{P}_1 \underline{\mathbf{X}}_i \simeq \mathbf{P}_j \underline{\mathbf{X}}_i \quad \text{for all } X_i \in \mathcal{X}$$

there is a non-trivial set of other cameras that see the same image

### Results

- importantly: If all calibration points  $X_i \in \mathcal{X}$  lie on a plane  $\varkappa$  then camera resection is non-unique and all image-equivalent camera centers lie on a spatial line  $\mathcal{C}$  with the  $C_\infty = \varkappa \cap \mathcal{C}$  excluded
- and more: by adding points  $X_i \in \mathcal{X}$  to  $\mathcal{C}$  we gain nothing
- there are additional image-equivalent configurations, see next



Case 4

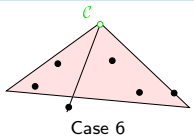
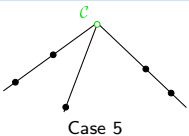
Proof sketch: If  $\mathbf{Q}, \mathbf{T}$  are suitable homographies then  $\mathbf{P}_1 \simeq \mathbf{Q}\mathbf{P}_0\mathbf{T}$ , where  $\mathbf{P}_0$  is canonical and the analysis can be made with  $\hat{\mathbf{P}}_j \simeq \mathbf{Q}^{-1}\mathbf{P}_j$

$$\mathbf{P}_0 \underbrace{\underline{\mathbf{T}}\underline{\mathbf{X}}_i}_{\underline{\mathbf{Y}}_i} \simeq \hat{\mathbf{P}}_j \underbrace{\underline{\mathbf{T}}\underline{\mathbf{X}}_i}_{\underline{\mathbf{Y}}_i} \quad \text{for all } Y_i \in \mathcal{Y}$$

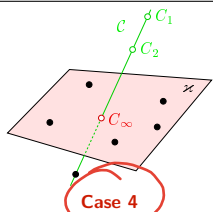
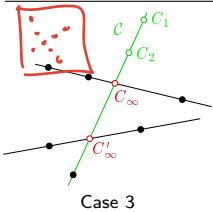
see [H&Z, Sec. 22.1.2] for a full prof



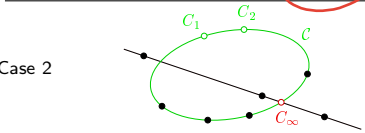
# cont'd (all cases)



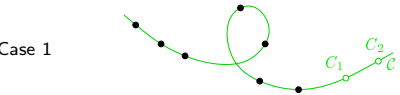
- points lie on three optical rays or one optical ray and one optical plane
- cameras  $C_1, C_2$  co-located at point  $C$
- Case 5: camera sees 3 isolated point images
- Case 6: cam. sees a line of points and an isolated point



- points lie on a line  $C$  and
  1. on two lines meeting  $C$  at  $C_\infty, C'_\infty$
  2. or on a plane meeting  $C$  at  $C_\infty$
- cameras lie on a line  $C \setminus \{C_\infty, C'_\infty\}$
- Case 3: camera sees 2 lines of points
- Case 4: **dangerous!**



- points lie on a planar conic  $C$  and an additional line meeting  $C$  at  $C_\infty$
- cameras lie on  $C \setminus \{C_\infty\}$  not necessarily an ellipse
- Case 2: camera sees 2 lines of points



- points and cameras all lie on a twisted cubic  $C$
- Case 1: camera sees points on a conic  
dangerous but unlikely to occur

## ► Three-Point Exterior Orientation Problem (P3P)

Calibrated camera rotation and translation from Perspective images of 3 reference Points.

**Problem:** Given  $\mathbf{K}$  and three corresponding pairs  $\{(m_i, X_i)\}_{i=1}^3$ , find  $\mathbf{R}$ ,  $\mathbf{C}$  by solving

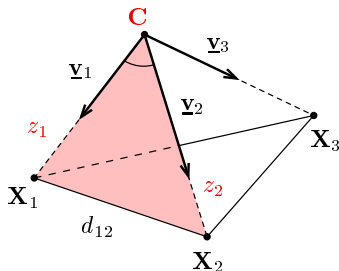
$X_i$ : finite

$$\lambda_i \underline{\mathbf{m}}_i = \mathbf{K} \mathbf{R} (\mathbf{X}_i - \mathbf{C}), \quad i = 1, 2, 3 \quad \mathbf{X}_i \text{ Cartesian}$$

1. Transform  $\underline{\mathbf{v}}_i \stackrel{\text{def}}{=} \mathbf{K}^{-1} \underline{\mathbf{m}}_i$ . Then

$$\lambda_i \underline{\mathbf{v}}_i = \mathbf{R} (\mathbf{X}_i - \mathbf{C}). \quad (10)$$

2. If there was no rotation in (10), the situation would look like this



3. and we could shoot 3 lines from the given points  $X_i$  in given directions  $\underline{\mathbf{v}}_i$  to get  $\mathbf{C}$
4. given  $\mathbf{C}$  we solve (10) for  $\lambda_i$ ,  $\mathbf{R}$

## ► P3P cont'd

### If there is rotation $\mathbf{R}$

1. Eliminate  $\mathbf{R}$  by taking

rotation preserves length:  $\|\mathbf{R}\mathbf{x}\| = \|\mathbf{x}\|$

$$|\lambda_i| \cdot \|\mathbf{v}_i\| = \|\mathbf{X}_i - \mathbf{C}\| \stackrel{\text{def}}{=} z_i \quad (11)$$

2. Consider only angles among  $\mathbf{v}_i$  and apply Cosine Law per triangle  $(\mathbf{C}, \mathbf{X}_i, \mathbf{X}_j)$   $i, j = 1, 2, 3, i \neq j$

$$d_{ij}^2 = z_i^2 + z_j^2 - 2 z_i z_j c_{ij},$$

$$z_i = \|\mathbf{X}_i - \mathbf{C}\|, \quad d_{ij} = \|\mathbf{X}_j - \mathbf{X}_i\|, \quad c_{ij} = \cos(\angle \mathbf{v}_i \mathbf{v}_j)$$

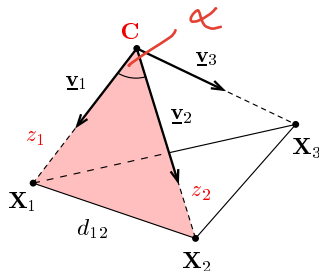
4. Solve the system of 3 quadratic eqs in 3 unknowns  $z_i$

[Fischler & Bolles, 1981]

there may be no real root

there are up to 4 solutions that cannot be ignored (verify on additional points)

5. Compute  $\mathbf{C}$  by trilateration (3-sphere intersection) from  $\mathbf{X}_i$  and  $z_i$ ; then  $\lambda_i$  from (11)
6. Compute  $\mathbf{R}$  from (10) we will solve this problem next  $\rightarrow 70$



Similar problems (P4P with unknown  $f$ ) at <http://aag.ciirc.cvut.cz/minimal/> (papers, code)

Thank You

