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

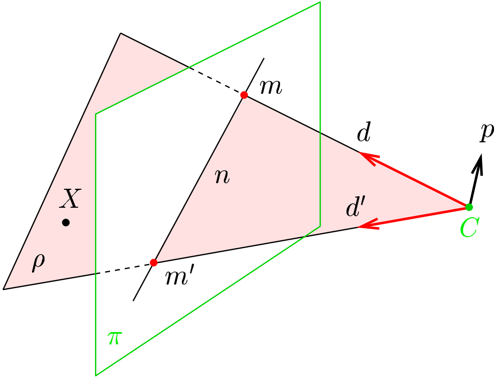
http://cmp.felk.cvut.cz
mailto:sara@cmp.felk.cvut.cz
phone ext. 7203

rev. October 10, 2023

Open Informatics Master's Course

▶Optical Plane

A spatial plane with normal p containing the projection center C and a given image line n.



optical ray given by m $\quad \mathbf{d} \simeq \mathbf{Q}^{-1}\underline{\mathbf{m}}$

optical ray given by m' $\quad \mathbf{d}' \simeq \mathbf{Q}^{-1}\underline{\mathbf{m}}'$

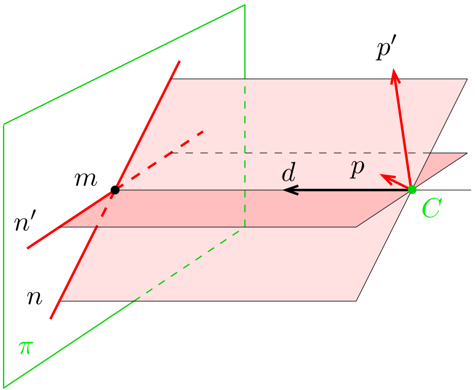
$$\mathbf{p} \simeq \mathbf{d} \times \mathbf{d}' = (\mathbf{Q}^{-1}\underline{\mathbf{m}}) \times (\mathbf{Q}^{-1}\underline{\mathbf{m}}') \overset{\circledast 1}{=} \mathbf{Q}^{\top}(\underline{\mathbf{m}} \times \underline{\mathbf{m}}') = \mathbf{Q}^{\top}\underline{\mathbf{n}}$$

- note the way $\mathbf{Q}$ factors out!

hence, $0 = \mathbf{p}^{\top}(\mathbf{X} - \mathbf{C}) = \underline{\mathbf{n}}^{\top} \underbrace{\mathbf{Q}(\mathbf{X} - \mathbf{C})}_{\to 30} = \underline{\mathbf{n}}^{\top}\mathbf{P}\underline{\mathbf{X}} = (\mathbf{P}^{\top}\underline{\mathbf{n}})^{\top}\underline{\mathbf{X}}$ for every X in plane ρ

optical plane is given by n: $\quad \underline{\boldsymbol{\rho}} \simeq \mathbf{P}^{\top}\underline{\mathbf{n}}$

$\underline{\boldsymbol{\rho}}$ are the plane's parameters: $\quad \rho_1\, x + \rho_2\, y + \rho_3\, z + \rho_4 = 0$



optical plane normal given by n is $\mathbf{p} = \mathbf{Q}^\top \underline{\mathbf{n}}$

optical plane normal given by n' is $\mathbf{p}' = \mathbf{Q}^\top \underline{\mathbf{n}}'$

The optical ray through their intersection is then

$$\mathbf{d} = \mathbf{p} \times \mathbf{p}' = (\mathbf{Q}^\top \underline{\mathbf{n}}) \times (\mathbf{Q}^\top \underline{\mathbf{n}}') = \mathbf{Q}^{-1}(\underline{\mathbf{n}} \times \underline{\mathbf{n}}') = \mathbf{Q}^{-1}\underline{\mathbf{m}}$$

General (finite) camera

$$\mathbf{P} = \begin{bmatrix}\mathbf{Q} & \mathbf{q}\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1^\top & q_{14} \\ \mathbf{q}_2^\top & q_{24} \\ \mathbf{q}_3^\top & q_{34}\end{bmatrix} = \mathbf{K}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix} = \mathbf{K}\,\mathbf{R}\begin{bmatrix}\mathbf{I} & -\mathbf{C}\end{bmatrix}$$

$\underline{\mathbf{C}} \simeq \mathrm{rnull}(\mathbf{P}),\quad \mathbf{C} = -\mathbf{Q}^{-1}\mathbf{q}$ — projection center (world coords.) →35

$\mathbf{d} = \mathbf{Q}^{-1}\,\underline{\mathbf{m}}$ — optical ray direction (world coords.) →36

$\mathbf{o} = \det(\mathbf{Q})\,\mathbf{q}_3$ — outward optical axis (world coords.) →37

$\underline{\mathbf{m}}_0 \simeq \mathbf{Q}\,\mathbf{q}_3$ — principal point (in image plane) →38

$\underline{\boldsymbol{\rho}} = \mathbf{P}^\top\,\underline{\mathbf{n}}$ — optical plane (world coords.) →39

$$\mathbf{K} = \begin{bmatrix} a\,f & -a\,f\,\cot\theta & u_0 \\ 0 & f/\sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$ camera (calibration) matrix (f, u_0, v_0 in pixels) →31

$\mathbf{R}$ — 3D rotation matrix (cam coords.) →30

$\mathbf{t}$ — 3D translation vector (cam coords.) →30

What Can We Do with An 'Uncalibrated' Perspective Camera?

How far is the engine from a given point on the tracks?

the distance between sleepers (ties) is 0.806m but we cannot count them, the image resolution is too low

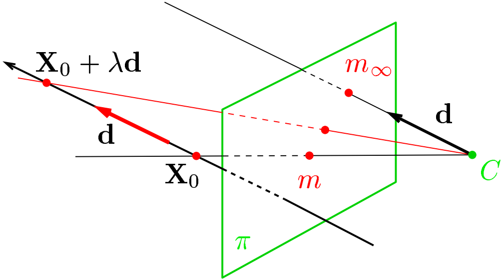
We will review some life-saving theory...
...and build a bit of geometric intuition...

In fact

- 'uncalibrated' = the image contains a 'calibrating object' that suffices for the task at hand

Vanishing point (V.P.): The limit m_∞ of the projection of a point $\mathbf{X}(\lambda)$ that moves along a space line $\mathbf{X}(\lambda) = \mathbf{X}_0 + \lambda\mathbf{d}$ infinitely in one direction.

the image of the point at infinity on the line

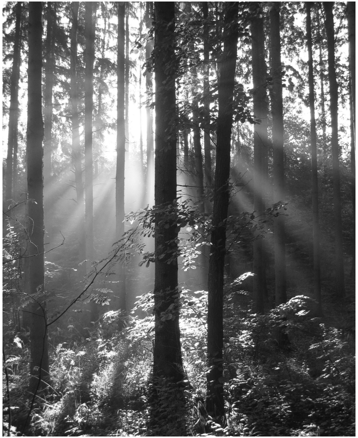


$$\underline{\mathbf{m}}_\infty \simeq \lim_{\lambda\to\pm\infty} \mathbf{P}\begin{bmatrix}\mathbf{X}_0 + \lambda\mathbf{d}\\ 1\end{bmatrix} = \cdots \simeq \mathbf{Q}\,\mathbf{d}$$

⊛ P1; 1pt: Prove (use Cartesian coordinates and L'Hôpital's rule)

- the V.P. of a spatial line with directional vector $\mathbf{d}$ is $\underline{\mathbf{m}}_\infty \simeq \mathbf{Q}\,\mathbf{d}$
- V.P. is independent on line position $\mathbf{X}_0$, it depends on its directional vector only
- all parallel (world) lines share the same (image) V.P., including the optical ray defined by m_∞

Some Vanishing Point "Applications"



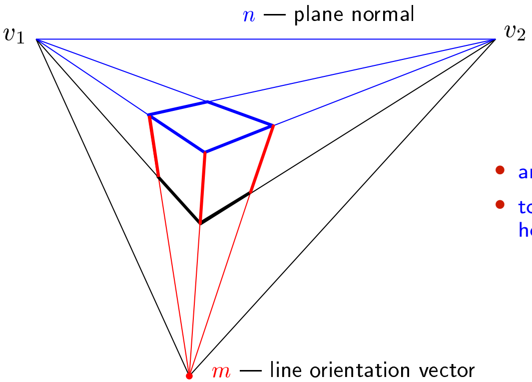
where is the sun?

what is the wind direction?
(must have video)

fly above the lane,
at constant altitude!

▶Vanishing Line

Vanishing line (V.L.): The set of vanishing points of all lines in a plane the image of the line at infinity in the plane and in all parallel planes (!)



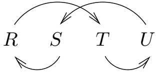
- any box with parallel edges
- top (blue) and bottom (black) box planes are parallel, hence they share V.L. n

- V.L. n corresponds to spatial plane of normal vector $\mathbf{p} = \mathbf{Q}^\top \underline{\mathbf{n}}$
  because this is the normal vector of a parallel optical plane (!) →39
- a spatial plane of normal vector $\mathbf{p}$ has a V.L. represented by $\underline{\mathbf{n}} = \mathbf{Q}^{-\top}\mathbf{p}$.

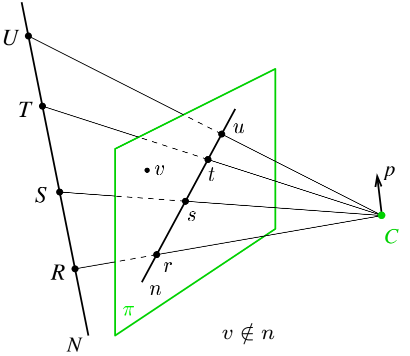
►Cross Ratio

Four distinct collinear spatial points R, S, T, U define cross-ratio

$$[RSTU] = \frac{|\overrightarrow{RT}|}{|\overrightarrow{SR}|}\,\frac{|\overrightarrow{US}|}{|\overrightarrow{TU}|}$$



a mnemonic (∞)



- $|\overrightarrow{RT}|$ – signed distance from R to T in the arrow direction
- each point X is once in numerator and once in denominator
- if X is 1st in a numerator term, it is 2nd in a denominator term
- there are six cross-ratios from four points:

$$[SRUT] = [RSTU],\ [RSUT] = \frac{1}{[RSTU]},\ [RTSU] = 1 - [RSTU],\ \cdots$$

Obs: $$[RSTU] = \frac{|\underline{\mathbf{r}}\ \underline{\mathbf{t}}\ \underline{\mathbf{v}}|}{|\underline{\mathbf{s}}\ \underline{\mathbf{r}}\ \underline{\mathbf{v}}|} \cdot \frac{|\underline{\mathbf{u}}\ \underline{\mathbf{s}}\ \underline{\mathbf{v}}|}{|\underline{\mathbf{t}}\ \underline{\mathbf{u}}\ \underline{\mathbf{v}}|}, \qquad |\underline{\mathbf{r}}\ \underline{\mathbf{t}}\ \underline{\mathbf{v}}| = \det\left[\underline{\mathbf{r}}\ \underline{\mathbf{t}}\ \underline{\mathbf{v}}\right] = (\underline{\mathbf{r}} \times \underline{\mathbf{t}})^\top \underline{\mathbf{v}} \tag{1}$$

mixed product

Corollaries:

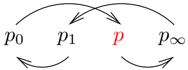
- cross ratio is invariant under homographies $\underline{\mathbf{x}}' \simeq \mathbf{H}\underline{\mathbf{x}}$ — proof: plug $\mathbf{H}\underline{\mathbf{x}}$ in (1): $(\mathbf{H}^{-\top}(\underline{\mathbf{r}} \times \underline{\mathbf{t}}))^\top \mathbf{H}\underline{\mathbf{v}}$
- cross ratio is invariant under perspective projection: $[RSTU] = [\,r\,s\,t\,u\,]$

- 4 collinear points: any perspective camera will “see” the same cross-ratio of their images
- we measure the same cross-ratio in image as on the world line
- one of the points R, S, T, U may be at infinity (we take the limit, in effect $\frac{\infty}{\infty} = 1$)

▸1D Projective Coordinates

The 1-D projective coordinate of a point P is defined by the following cross-ratio:

$$[P] = [P_0\, P_1\, P\, P_\infty] = [p_0\, p_1\, p\, p_\infty] = \frac{|\overrightarrow{p_0\, p}|}{|\overrightarrow{p_1\, p_0}|}\,\frac{|\overrightarrow{p_\infty\, p_1}|}{|\overrightarrow{p\, p_\infty}|} = [p]$$



naming convention:

P_0 – the origin $\quad [P_0] = 0$

P_1 – the unit point $\quad [P_1] = 1$

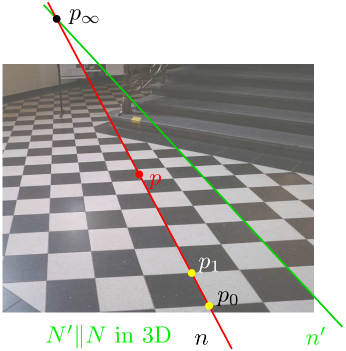
P_∞ – the supporting point $\quad [P_\infty] = \pm\infty$

$$[P] = [p]$$

$[P]$ is equal to Euclidean coordinate along N

$[p]$ is its measurement in the image plane

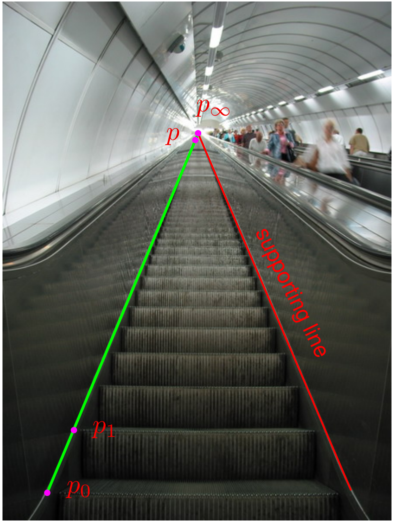
if the sign is not of interest, any cross-ratio containing $|p_0\, p|$ does the job



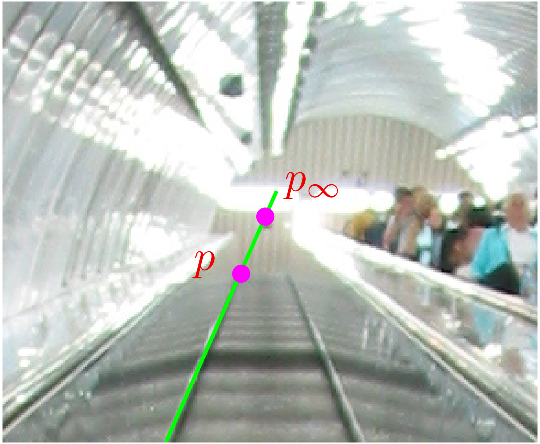
Applications

- Given the image of a 3D line N, the origin, the unit point, and the vanishing point, then the Euclidean coordinate of any point $P \in N$ can be determined →48
- Finding V.P. of a line through a regular object →49

Application: Counting Steps



- Namesti Miru underground station in Prague

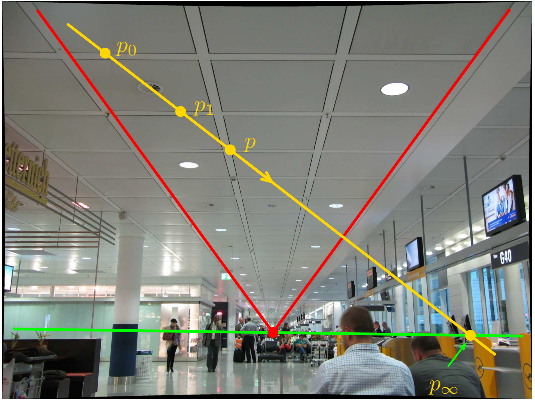


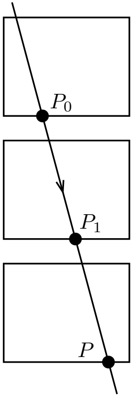
detail around the vanishing point (w/ strong aliasing)

Result: $[P] = 214$ steps (correct answer is 216 steps)

4Mpx camera

Application: Finding the Horizon from Repetitions





in 3D: $|P_0 P| = 2|P_0 P_1|$ then [H&Z, p. 218]

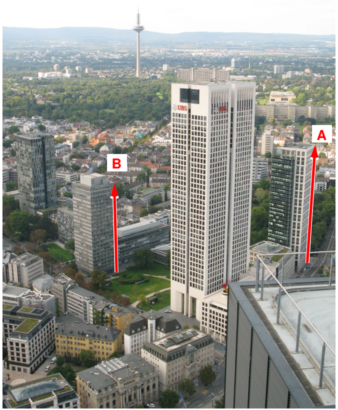
$$[P_0 P_1 P P_\infty] = \frac{|P_0 P|}{|P_1 P_0|} = 2 \quad \Rightarrow \quad x_\infty = \frac{x_0\,(2\,x - x_1) - x\,x_1}{x + x_0 - 2\,x_1}$$

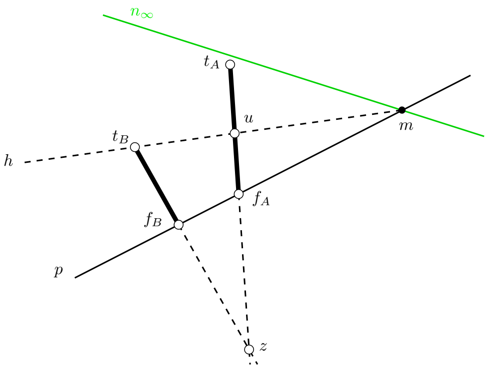
- x – 1D coordinate along the yellow line, positive in the arrow direction
- could be applied to counting steps (→48) if there was no supporting line

⊛ P1; 1pt: How high is the camera above the floor?

⊛ 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

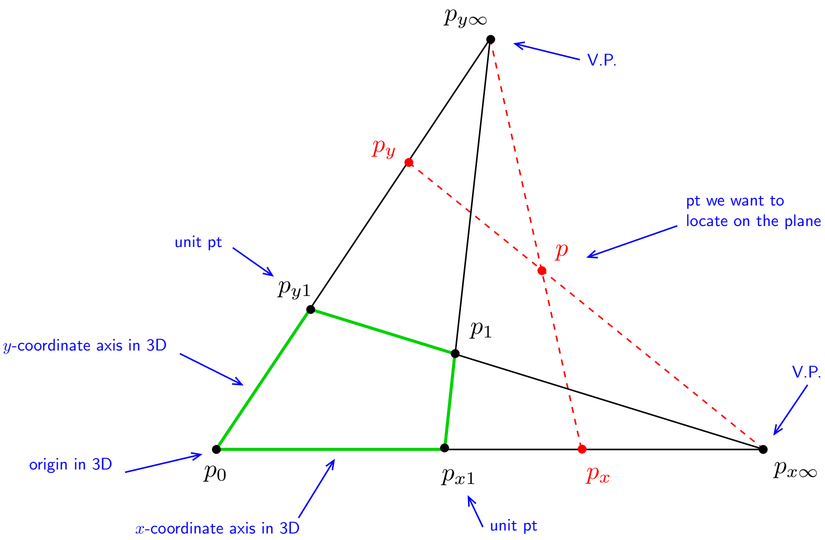




Hints

1. What are the interesting properties of line h connecting the top t_B of Buiding B with the point m at which the horizon intersects the line p joining the foots f_A, f_B of both buildings? [1 point]
2. How do we actually get the horizon n_∞? (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



$$[P_x] = [P_0 \; P_{x1} \; P_x \; P_{x\infty}] \qquad [P_y] = [P_0 \; P_{y1} \; P_y \; P_{y\infty}]$$

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)

Module III

Computing with a Single Camera

covered by

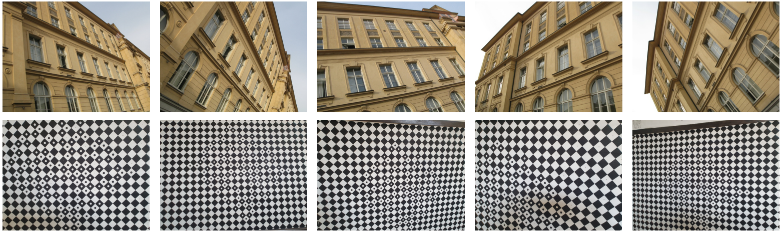
[1] [H&Z] Secs: 8.6, 7.1, 22.1

[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

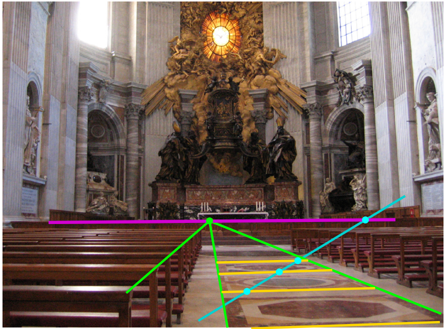
[3] [Golub & van Loan 2013, Sec. 2.5]

Obtaining Vanishing Points and Lines

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

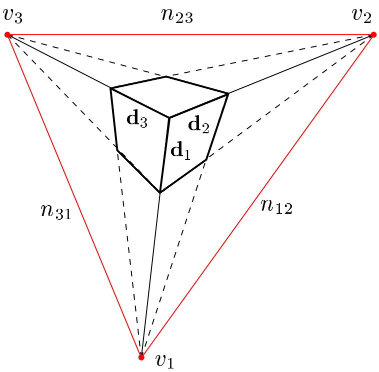


- vanishing line can be obtained from vanishing points and/or regularities (→49)



▶Camera Calibration from Vanishing Points and Lines

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



$$\mathbf{d}_i = \lambda_i \mathbf{Q}^{-1} \underline{\mathbf{v}}_i, \qquad i = 1,2,3 \quad \to 43$$

$$\mathbf{p}_{ij} = \mu_{ij} \mathbf{Q}^\top \underline{\mathbf{n}}_{ij}, \quad i,j = 1,2,3,\ i \neq j \quad \to 39 \qquad (2)$$

- method: eliminate λ_i, μ_{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

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

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

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

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

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

- n_{ij} may be constructed from non-orthogonal v_i and v_j, e.g. using the cross-ratio
- $\boldsymbol{\omega}$ is a homogeneous, symmetric, definite 3×3 matrix (5 DoF) — IAC = Image of Absolute Conic
- equations are quadratic in $\mathbf{K}$ but linear in $\boldsymbol{\omega}$

►cont'd

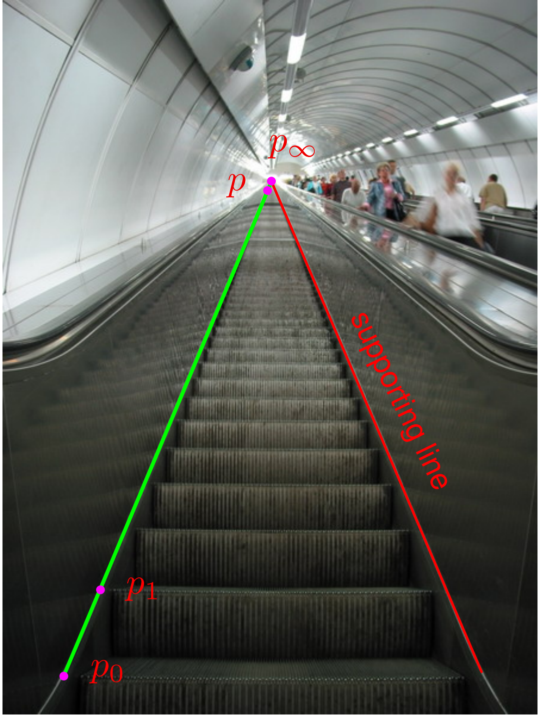
	configuration	equation	# constraints
(3)	orthogonal vanishing points	$\underline{\mathbf{v}}_i^\top \boldsymbol{\omega}\, \underline{\mathbf{v}}_j = 0$	1
(4)	orthogonal vanishing lines	$\underline{\mathbf{n}}_{ij}^\top \boldsymbol{\omega}^{-1} \underline{\mathbf{n}}_{ik} = 0$	1
(5)	vanishing points orthogonal to vanishing lines	$\underline{\mathbf{n}}_{ij} = \varkappa\, \boldsymbol{\omega}\, \underline{\mathbf{v}}_k$	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

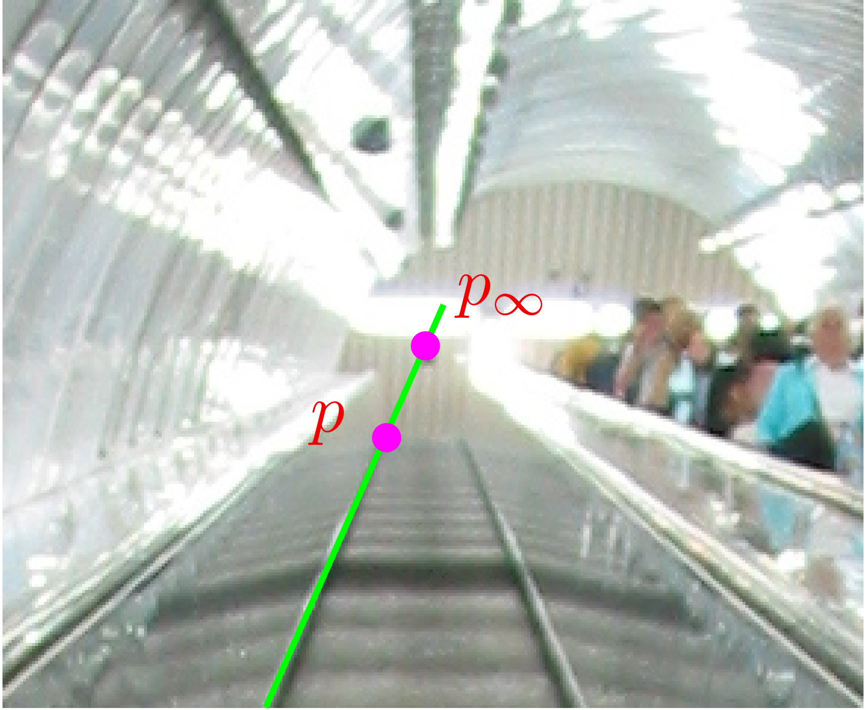
- These are homogeneous linear equations for the 5 parameters in $\boldsymbol{\omega}$ or $\boldsymbol{\omega}^{-1}$ — $\varkappa$ can be eliminated from (5)
- When $\mathbf{w} = \mathrm{vec}(\boldsymbol{\omega}) \in \mathbb{R}^6$, it has the form of $\mathbf{D}\mathbf{w} = \mathbf{0}$, $\mathbf{D} \in \mathbb{R}^{k \times 5}$
- With $k = 5$ constraints, we have $\mathrm{rank}(\mathbf{D}) = 5$, hence there is a unique solution for the homogeneous $\mathbf{w}$.
- We get $\mathbf{K}$ from $\boldsymbol{\omega}^{-1} = \mathbf{K}\mathbf{K}^\top$ by Choleski decomposition

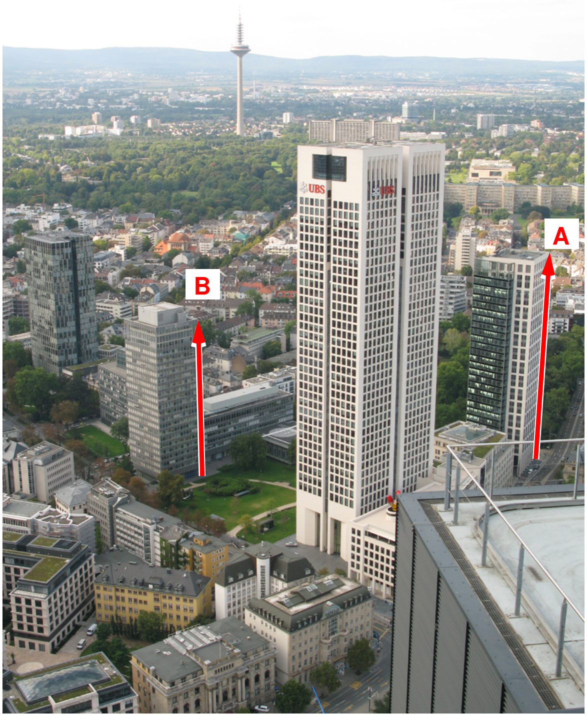
the decomposition returns a positive definite upper triangular matrix

one avoids solving an explicit set of quadratic equations for the parameters in $\mathbf{K}$

Thank You


p_∞
p
supporting line
p_1
p_0


p_∞
p


A
B