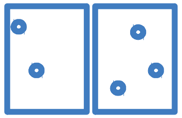
Lecture 4.1
Two View Geometry

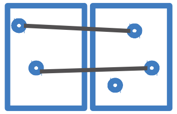
ENV408: Optical Sensing & Modeling for Earth Observations

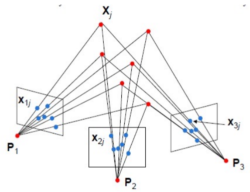
Jan Skaloud

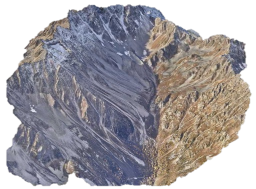
Flight planning → Image acquisition → (Trajectory e.g. navigation sensors) → Images

Feature detection → Feature matching → Orientation optimization → Production (e.g. surface, ortho-photo)









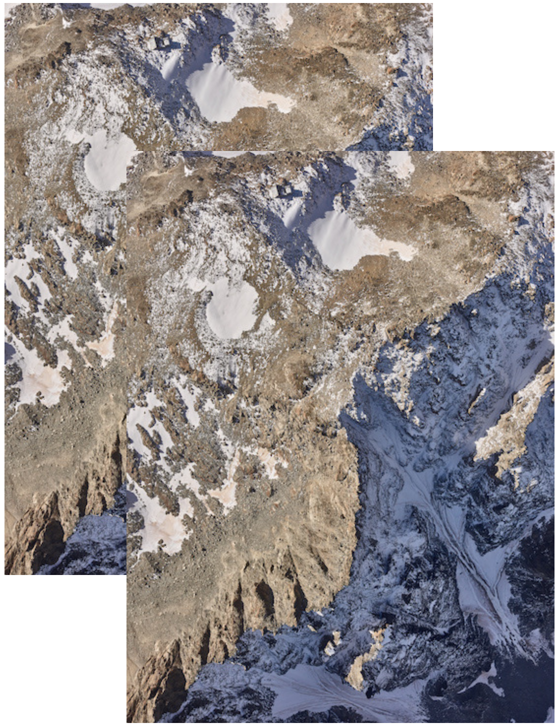
Lectures

- Image primes (L1)
- Salient *features* (L2)
- Image *orientation* (L3)
- *Optimization* (calibration) (L4)
- Mapping products (L5)

Exercises

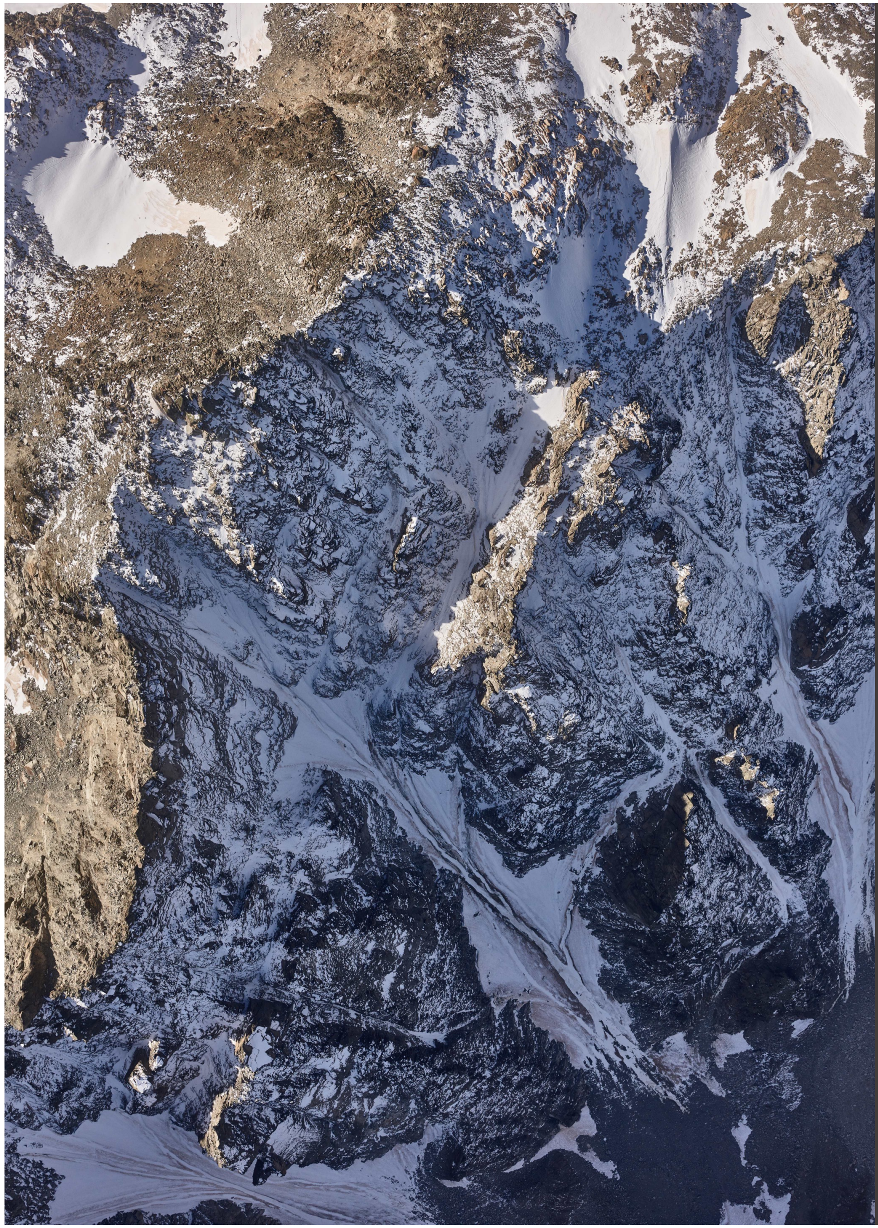
- Image ‘corrections’ (Lab01)
- Detection & matching (Lab02)
- Approx. absolute orientation (Lab03)
- Approx. relative orientation (Lab04)
- Calibration, DEM, ortho-photo (Lab05)

Relative orientation



Today
Understand how to reconstruct simultaneously 3D scene structure and camera pose from multiple images

Tomorrow
Use undistorted (Lab01) matched key-points (Lab02) to orient two images and triangulate key-point coordinates in 3D.



Epipolar Geometry

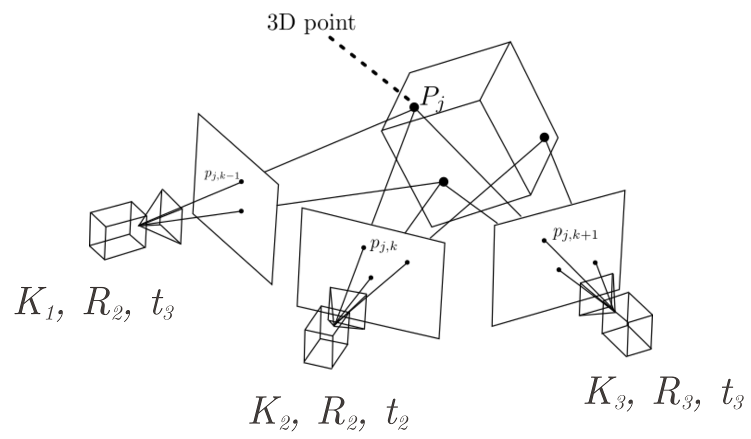
Essential and Fundamental Matrices

8 points algorithm

Two views geometry - recapitulation

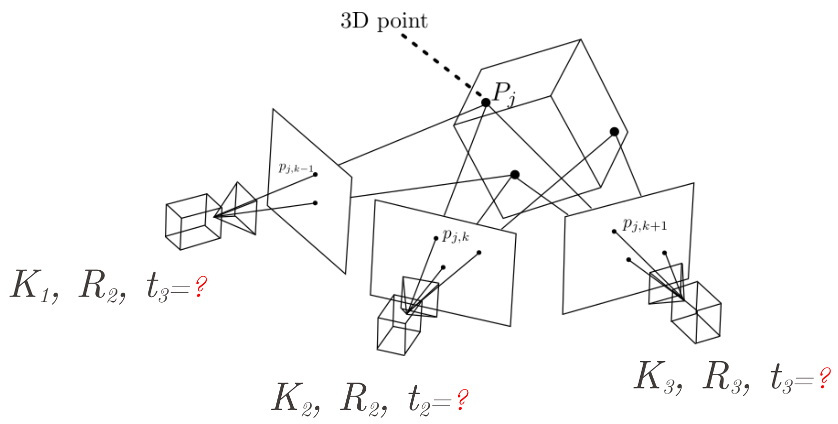
Last week: depth from stereo vision
(= 3D reconstruction)

- Assumptions: **known** $K_i,\ R_i,\ t_i$ (i.e. calibrated & oriented camera)
- Goal: recover the 3D structure from images



This week: Structure from motion
(SFM)

- Assumptions: **unknown** $K_i,\ R_i,\ t_i$
- Goal: recover simultaneously 3D structure + camera pose



Correspondence problem

Triangulation prerequisits

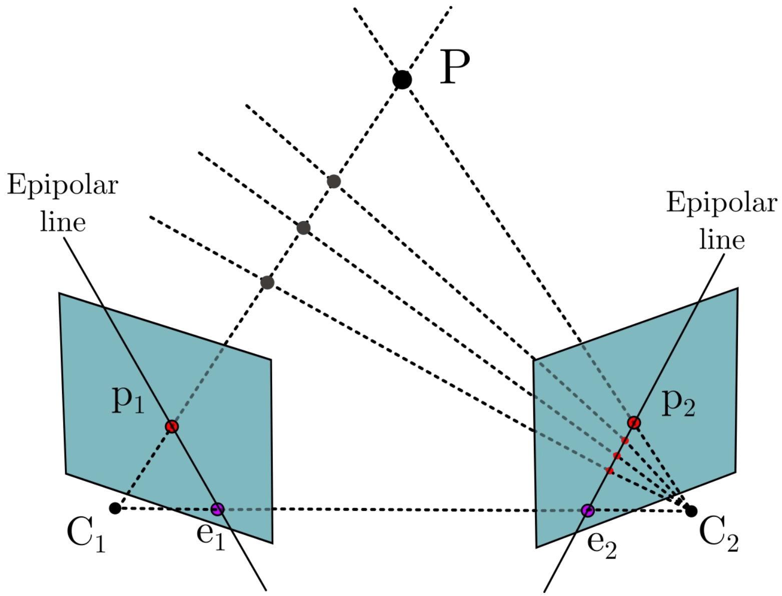
- Pose (R and t) is known (at least relatively)
- Image correspondences exist for a set of points $P_i \quad i=1 \ldots n$

Questions

- Given a point on the left-image, p_L, where is its correspondence, p_R, on the right image?
- Note: 2D exhaustive search is very expensive (computationally)

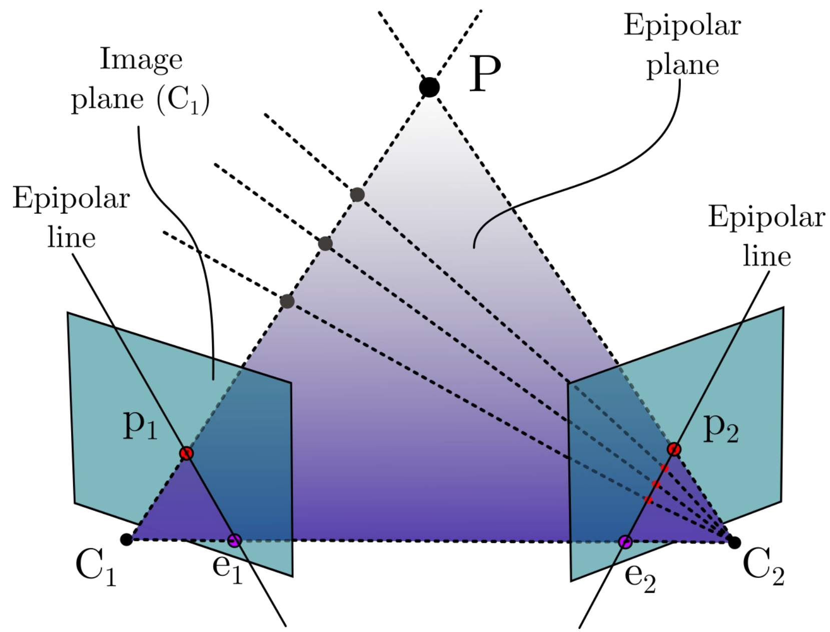
Answer:

Potential matches have to lie on an epipolar line! (see after)



Epipolar Geometry

- **Epipolar plane**: 3D plane formed by $\mathbf{C_1}, \mathbf{C_2}$ (cam. centers) & $\mathbf{P}$
- **Epipoles $\mathbf{e_1}, \mathbf{e_2}$:** intersection of the line $\mathrm{C_1}, \mathrm{C_2}$ with image planes
- **Epipolar line:** Intersection of epipolar plane with image plane
- **Epipolar constraint**: given P, corresponding points $\mathrm{p1}, \mathrm{p2}$ **must** lie on their respective epipolar lines



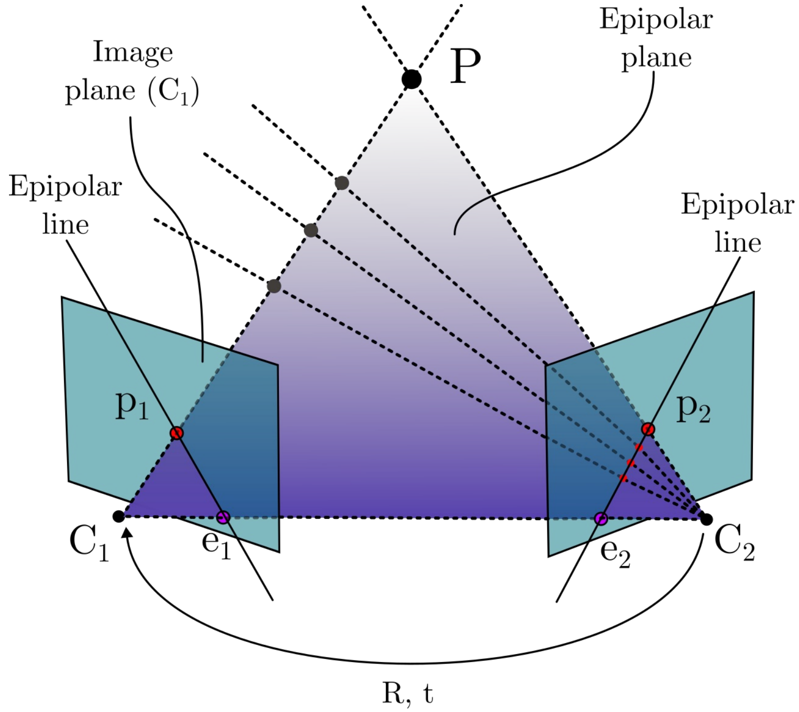
Epipolar Constraint 1/3

Formulation via epipolar lines

- $\mathbf{R}, \mathbf{t}$: rotation and translation relative from C2 to C1
- Point P in:
  - Camera 1 frame $\overrightarrow{C1, P} = P_1 = \mu_1 p_1$
  - Camera 2 frame $\overrightarrow{C2, P} = P_2 = \mu_2 p_2$

$$\Downarrow$$

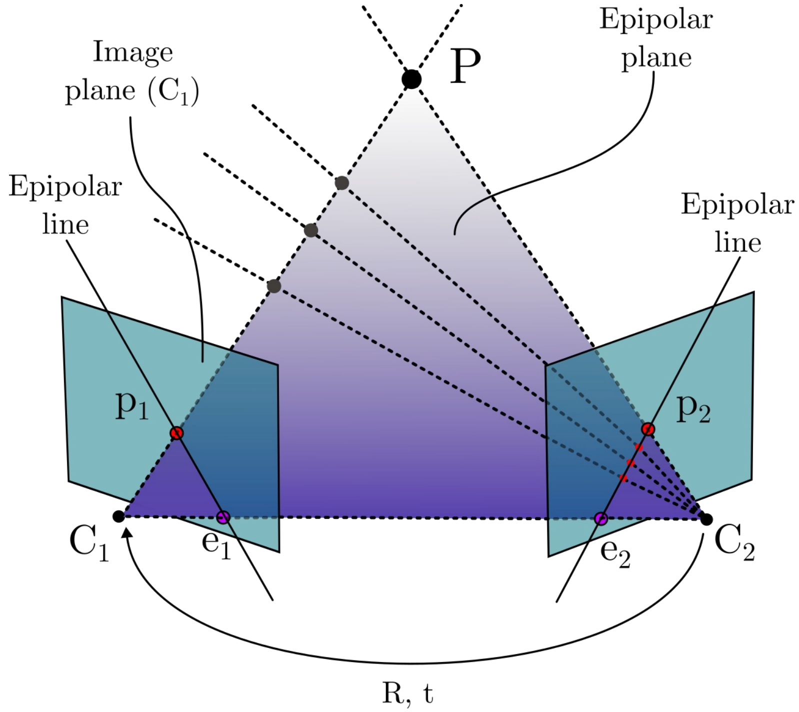
$$\mu_1 p_1 = R \mu_2 p_2 + t$$



Epipolar Constraint 2/3

Formulation via epipolar lines

- The 3 vectors $\overrightarrow{C_1, P_1}$, $\overrightarrow{C_2, P_2}$ and $\overrightarrow{C_1, C_2} = t$ must be coplanar
- Mathematically equivalent to null vector triple product $a \cdot (b \times c) = 0$
- This gives: $P_1 \cdot (t \times \underbrace{RP_2}_{\text{P}_2 \text{ in camera 1 frame}}) = 0$



Epipolar Constraint 3/3

Reminder: vector *cross* product equivalent to skew-symmetric matrix multiplication $[t_\times]$

$$P_1 \cdot (t \times RP_2) = 0$$

$$\Updownarrow$$

$$P_1^T [t_\times] R P_2 = 0$$

$$\Updownarrow$$

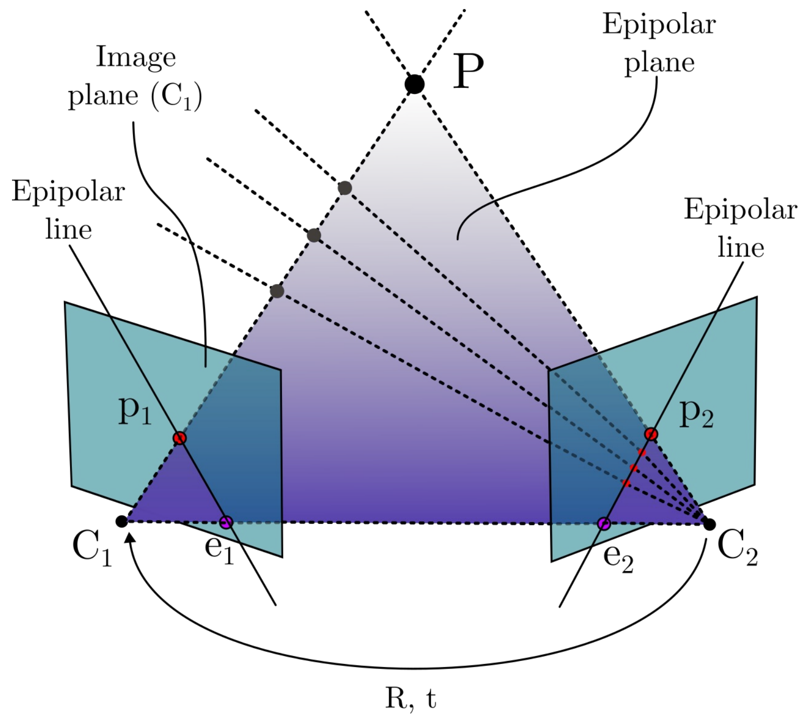
Only vector direction counts

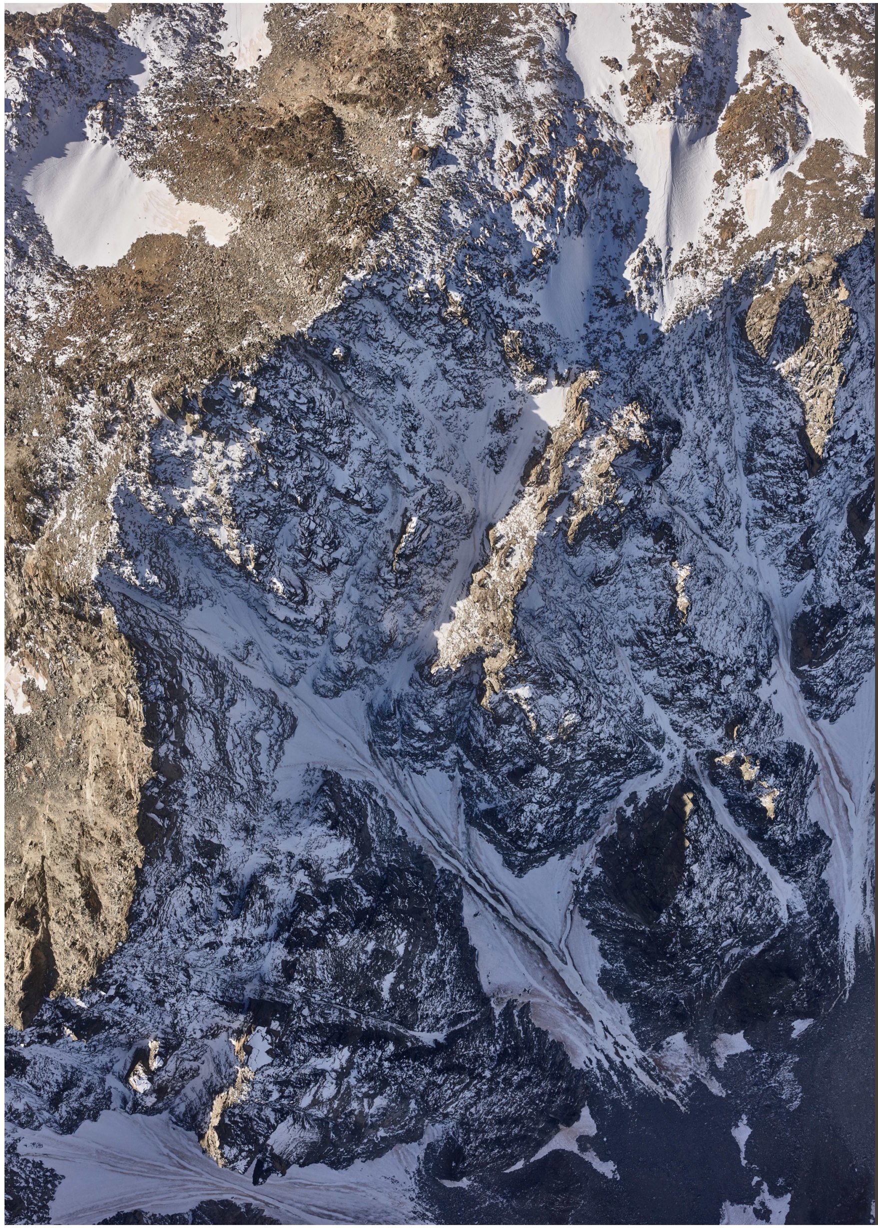
$$p_1^T [t_\times] R\, p_2 = 0$$

Defining Essential matrix $\mathbf{E}$:

$$E \equiv [t_\times] R$$

$$p_2^T E\, p_1 = 0$$





Epipolar Geometry

Essential and Fundamental Matrices

8 points algorithm

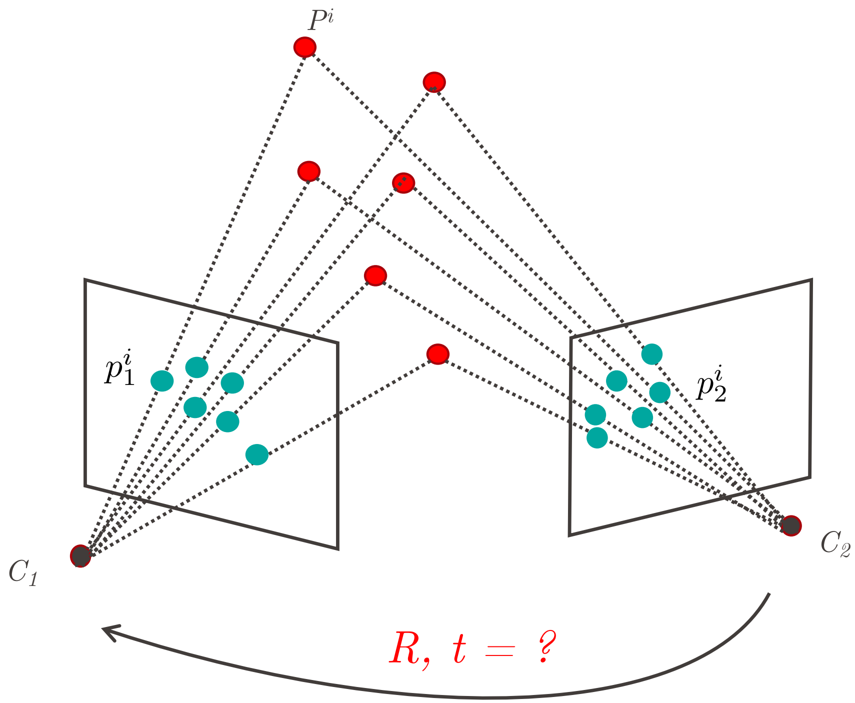
Relative orientation (SFM)

Given a set of $i=(1..n)$ point correspondences $p_1^i = (u_1^i,\, v_1^i)^T,\; p_2^i = (u_2^i,\, v_2^i)^T$ for 2 images, estimate simultaneously:

- The 3D points P^i
- The camera relative-orientation/pose $(R,\, t)$
- Camera intrinsic K1, K2, satisfying:

$$\mu_1^i \begin{pmatrix} u_1^i \\ v_1^i \\ 1 \end{pmatrix} = K_1[I\,|\,0] \cdot \begin{pmatrix} X^i \\ Y^i \\ Z^i \\ 1 \end{pmatrix}$$

$$\mu_2^i \begin{pmatrix} u_2^i \\ v_2^i \\ 1 \end{pmatrix} = K_2[R\,|\,t] \cdot \begin{pmatrix} X^i \\ Y^i \\ Z^i \\ 1 \end{pmatrix}$$



Epipolar geometry – calibrated camera

- Normalized undistorted coordinates

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = K_1^{-1} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = K_2^{-1} \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \qquad K = \begin{pmatrix} c & 0 & x_{PPS} + c_x \\ 0 & c & y_{PPS} + c_y \\ 0 & 0 & 1 \end{pmatrix}$$

$x_2^T E x_1 = 0$ (notation polycopié)

$p_2^T E p_1 = 0$ (notation slides + labs, same thing)

$$\begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}^T E \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = 0$$

Essential matrix $E = [t_\times]R$

-

Epipolar geometry – uncalibrated camera

Without the knowledge of $\mathbf{K}$: p_i can only be defined by u,v since x,y are unknown

$$\begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix}^T E \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = 0$$

$$\Downarrow$$

$$\begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix}^T \boxed{K_2^{-T} E K_1^{-1}} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = 0$$

Fundamental matrix $F = (K_2^T)^{-1} E K_1^{-1} \implies \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix}^T \boxed{F} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = 0$

Epipolar geometry – system of equations

Each pair of point correspondences $p_1 = (u_1, v_1, 1)^T, \quad p_2 = (u_2, v_2, 1)^T$ provides a linear equation:

$$p_2^T E p_1 = 0 \qquad E = \begin{pmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{pmatrix}$$

$$\Downarrow$$

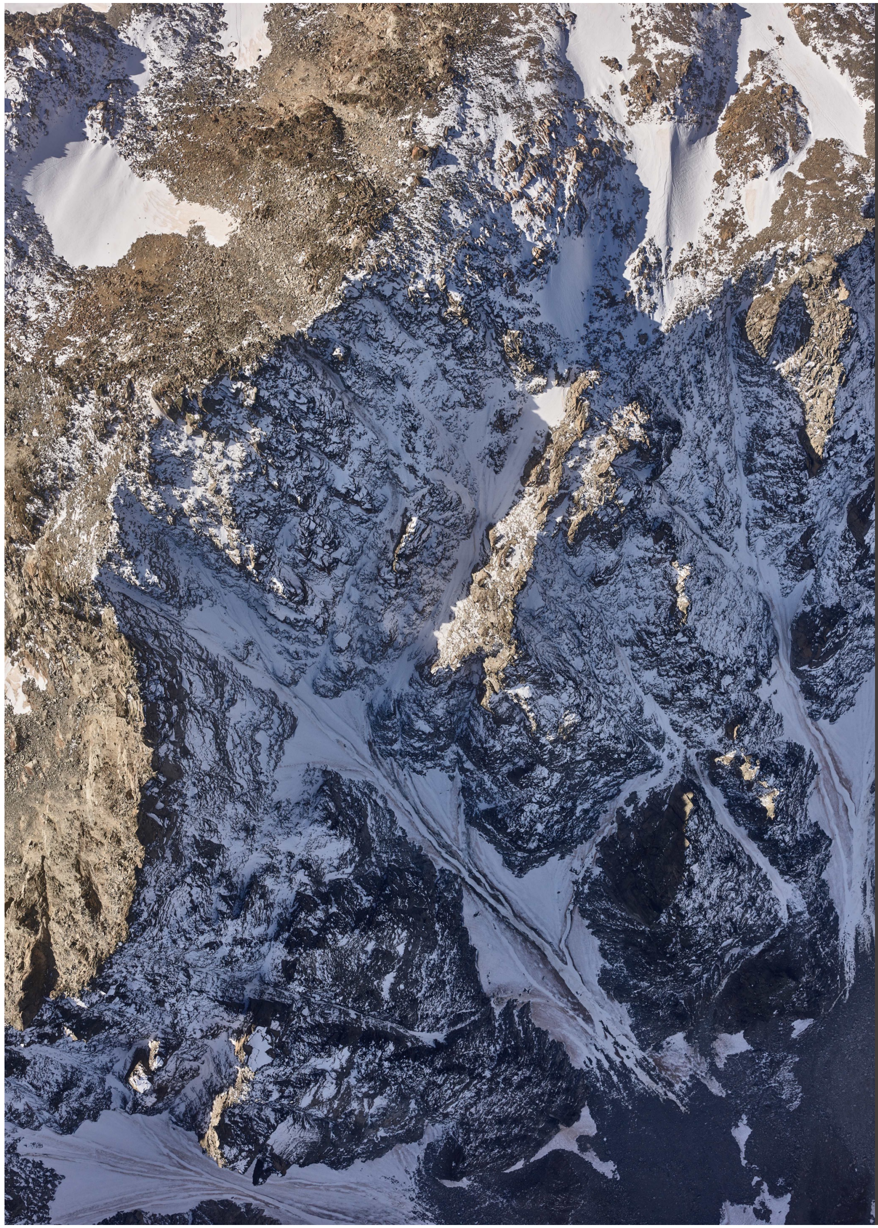
$$u_2 u_1 e_{11} + u_2 v_1 e_{12} + u_2 e_{13} + v_2 u_1 e_{21} + v_2 v_1 e_{22} + v_2 e_{23} + u_1 e_{31} + v_1 e_{32} + e_{33} = 0$$

Given enough correspondences, E (or F) can be obtained

1. What is the minimum number of correspondences ?
2. Can R, t can be recovered from E ?
3. (In more general case, can R, t, K_1, K_2 be recovered from F ?)

Epipolar geometry – inverse problem for E

- How many knowns per n ?
  - per correspondence:
  - per n :
- How many unknowns per n ?
  - per correspondence:
  - general:
  - together:
- When a solution exist?



Epipolar Geometry

Essential and Fundamental Matrices

8 points algorithm

Historical development

Kruppa – Determined the min. no. of correspondences (five), 11 solutions

Demazure – Showed that there is at most 10 distinct solutions

Nister – 1st efficient and non iterative solution (basis decomposition)**

1913

1981

1988

1996

2004

Longuet-Higgins – Easy implementation, 8-point algorithm (NASA-rover)*

Philipp – Described and iterative algorithm to find the solutions

* H. Christopher Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, Nature, 1981

**D. Nister, An Efficient Solution to the Five-Point Relative Pose Problem, PAMI, 2004.

The 8-point algorithm – formation of constraints

- For 1 point, we have from $p_2^T E p_1 = 0$

$$u_2 u_1 e_{11} + u_2 v_1 e_{12} + u_2 e_{13} + v_2 u_1 e_{21} + v_2 v_1 e_{22} + v_2 e_{23} + u_1 e_{31} + v_1 e_{32} + e_{33} = 0$$

- For n points (when omitting bars)

$$\underbrace{\begin{pmatrix} u_2^1 u_1^1 & u_2^1 v_1^1 & u_2^1 & v_2^1 u_1^1 & v_2^1 v_1^1 & v_2^1 & u_1^1 & v_1^1 & 1 \\ u_2^2 u_1^2 & u_2^2 v_1^2 & u_2^2 & v_2^2 u_1^2 & v_2^2 v_1^2 & v_2^2 & u_1^2 & v_1^2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_2^n u_1^n & u_2^n v_1^n & u_2^n & v_2^n u_1^n & v_2^n v_1^n & v_2^n & u_1^n & v_1^n & 1 \end{pmatrix}}_{Q \text{ (known)}} \underbrace{\begin{pmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{31} \\ e_{32} \\ e_{33} \end{pmatrix}}_{E_s \text{ (stacked } E \text{ - unknown)}} = 0 \quad \Longrightarrow \quad Q \cdot E_s = 0$$

Q (known)

E_s (stacked E - unknown)

The 8-point algorithm – finding E

Minimum solution $Q \cdot E_s = 0$

- $Q_{(n \times 9)}$ - a unique (up to a scale) solution is possible if matrix rank = …. ?
- Each correspondence gives 1 independent equation.
- Hence, … correspondences (non-planar) needed

Over-determined solution ($n >$?)

- By minimizing $||Q \cdot E_s||^2 = E_s^T Q^T Q E_s$ subject to constraint $||E_s||^2 = 1$
- Solution E_s is an eigenvector corresponding to the smallest eigen value of Q
- Via SVD of $Q^T Q$ matrix that is in this case equivalent to SVD of Q*

** K. Inkilä, 2005, Homogeneous least square problem, Photogrammetric Journal of Finland.*

The 8-point algorithm – SVD of Q in Python


```
Q = np.zeros((num_points,9))
    for i in range(num_points):
        Q[i,:] = np.kron( p1[:,i], p2[:,i] ).T

_, _, Vt= np.linalg.svd(Q, full_matrices = False)
     E = np.reshape(Vt[-1,:], (3,3)).T
```

Extracting R, t from E

I. Enforcing E to be in the "E-space"

- Singular value decomposition $E = U \Sigma V^T$
- "In case of no-errors", perfect correspondences: $\Sigma = diag\left(\sigma, \sigma, 0\right)$
- Due to errors: $\tilde{E} = U\, diag\left(\sigma_1, \sigma_2, \sigma_3\right) V^T, \quad \sigma_1 \geq \sigma_2 \geq \sigma_3$
- Choosing $\hat{E} = U\, diag\left(\sigma, \sigma, 0\right) V^T, \quad \sigma = (\sigma_1 + \sigma_2)/2$
- … satisfies E-space, but there could be another E leading to a smaller $||Q \cdot E_s||^2$
- Python

```
# Enforce det(E)=0 by projecting E on a set of 3x3 orthogonal matrices
    U, S, Vt = np.linalg.svd(E)
    S[0] = s[1] = (s[0]+s[1])/2
    S[2] = 0
    Ehat = U @ np.diag(S) @ Vt
```

Extracting R, t from E

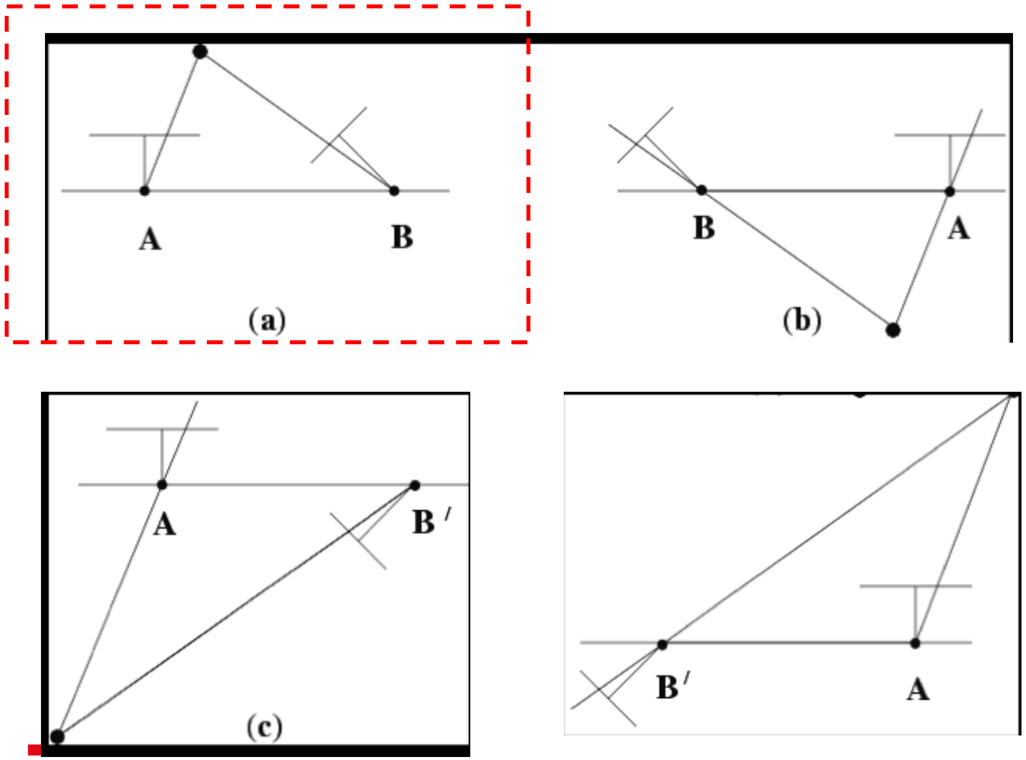
II. Finding t

- $RR^T{=}1$, thus: $EE = [t_\times]RR^T[t_\times]^T = [t_\times][t_\times]^T = [t_\times][-t_\times] = -[t_\times]^2$
- Reminder: $[t_\times] = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$ and $||t||_2 = 1$ (scale not recovered from E)
- Thus $-[t_\times]^2 = \begin{pmatrix} -t_z^2 - t_y^2 & t_x t_y & t_x t_z \\ t_x t_y & -t_z^2 - t_x^2 & t_z t_z \\ t_x t_z & t_y t_z & -t_y^2 - t_x^2 \end{pmatrix}$
- Since $||t||_2 = 1$, we obtain a matrix, from which diagonal we can obtain the absolute entries of t

$$-[t_\times]^2 = \begin{pmatrix} 1 - t_x^2 & t_x t_y & t_x t_z \\ t_x t_y & 1 - t_y^2 & t_z t_z \\ t_x t_z & t_y t_z & 1 - t_z^2 \end{pmatrix}$$

4 possible solutions for R, t

- However, the only plausible solution is the one when P lies in front-view of both cameras



- The are 4 possibilities to test”

$$\hat{R} = U \begin{pmatrix} 0 & \mp 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} V^T$$

$$[\hat{t}_\times] = U \begin{pmatrix} 0 & \mp 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Sigma U^T$$

$$[\hat{t}_\times] = \begin{pmatrix} 0 & -\hat{t}_z & \hat{t}_y \\ \hat{t}_z & 0 & -\hat{t}_x \\ -\hat{t}_y & \hat{t}_x & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} \hat{t}_x \\ \hat{t}_y \\ \hat{t}_z \end{pmatrix}$$

Remaining problem:

- Can R, t, K_1, K_2 be recovered from F ?

-

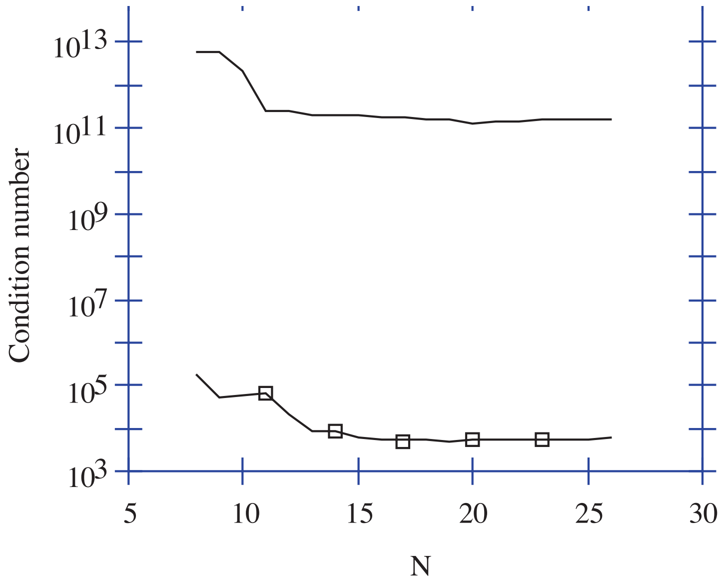
Practical issue correction

"Noise" in data

- E matrix near singular – points lying on the same 2D plane, small parallax (disparity)

Solution

- Translate all image points coordinates to a centroid
- Scale them so that the average distance from center is sqrt(2), i.e. $p_i = (1,\ 1,\ 1)$
- Improves condition number – solution stability!



Hartley, R.I., 2012: In defense of the 8-point algorithm. *IEEE Trans. Pattern Analysis*, 19(6), 580-593

Understanding - self assessment

1. What is the minimum number of correspondences between 2 images to perform relative orientation with a calibrated camera &why?
2. Can you provide a geometrical interpretation of the epipolar constraint?
3. How would you derive the epipolar constraint?
4. How is the essential matrix defined?
5. What are the properties of the essential matrix?
6. How are the essential and fundamental matrices related?
7. Is it always important to normalize the point coordinates in the 8-point algorithm?
8. How is the normalization of image coordinates performed?
9. How is the quality of the derived translation and rotation measured?