3D Vision

Marc Pollefeys
Daniel Barath

Spring 2024

ETH

Schedule

Feb 19	Introduction
Feb 26	Geometry, Camera Model, Calibration
Mar 4	Guest lecture + Features, Tracking / Matching
Mar 11	**Project Proposals by Students**
Mar 18	3DV conference
Mar 25	**Structure from Motion (SfM) + papers**
Apr 1	Easter break
Apr 8	Dense Correspondence (stereo / optical flow) + papers
Apr 15	Bundle Adjustment & SLAM + papers
Apr 22	**Student Midterm Presentations**
Apr 29	Multi-View Stereo & Volumetric Modeling + papers
May 6	3D Modeling with Depth Sensors + papers
May 13	Guest lecture + papers
May 20	Holiday

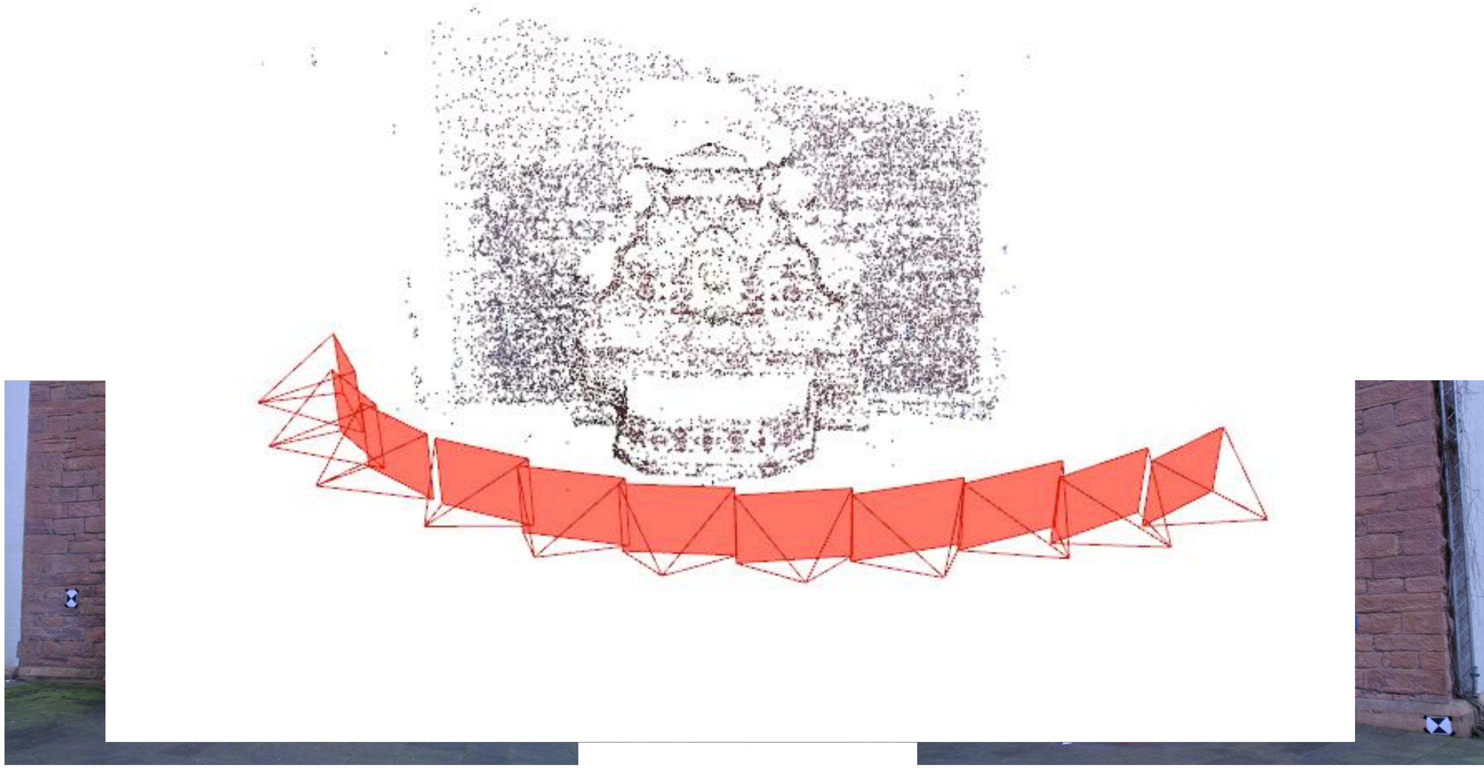
3D Vision – Class 4

Structure-from-Motion

Chapter 7 in Szeliski's Book
Chapter 9 in Hartley & Zisserman (online)
Tutorial chapters 3.2 and 4

ETH

Last Lecture: Local Features



Features are key component of many 3D Vision algorithms

ETH

Today: Structure-from-Motion (SfM)

Rome dataset

74,394 images

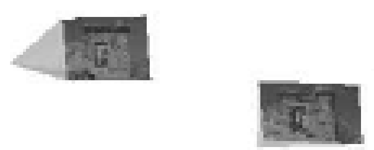
Johannes L. Schönberger and Jan-Michael Frahm. Structure-from-Motion revisited, CVPR 2016

ETH

Topics Today

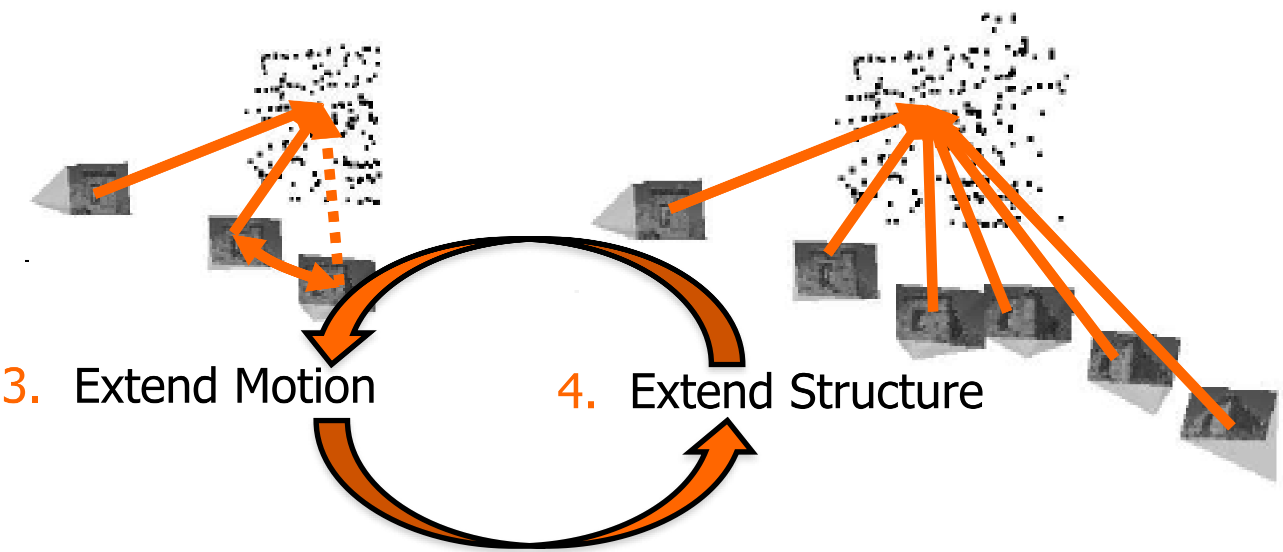
- Estimate motion between two images
  - Epipolar geometry
  - Two view Structure-from-Motion
- Estimate structure from motion
  - Triangulation
- Estimate camera pose from structure
  - Absolute camera pose solvers

Sequential / Incremental SfM



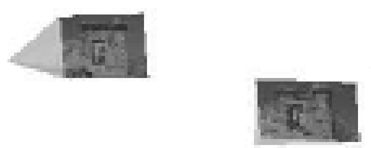
1. Initialize Motion

2. Initialize Structure

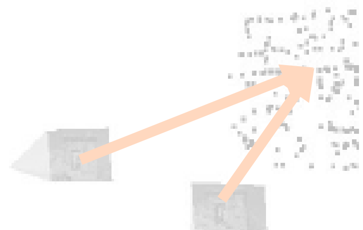


3. Extend Motion
4. Extend Structure

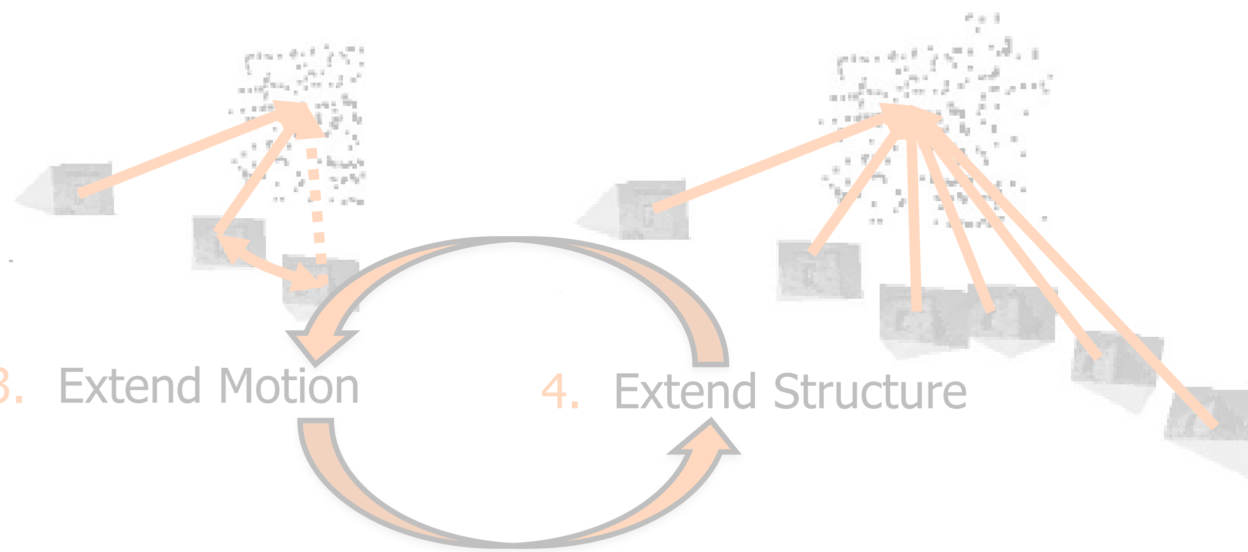
Sequential / Incremental SfM



1. **Initialize Motion**

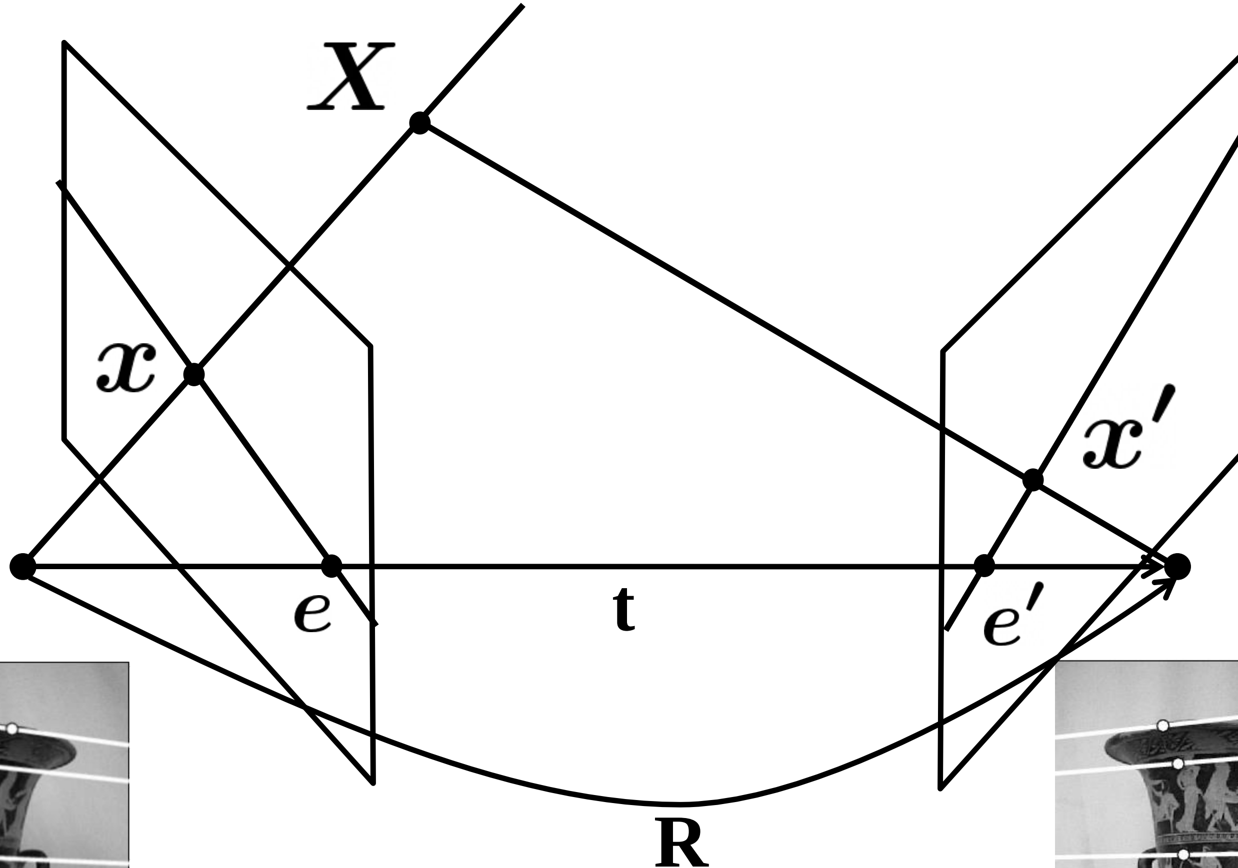


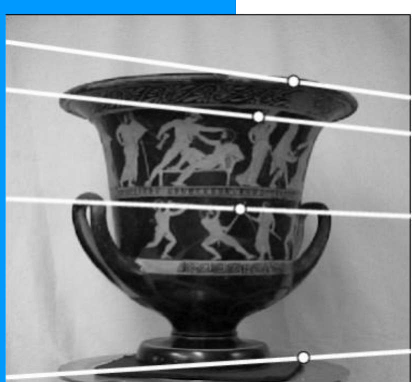
2. Initialize Structure



- **Two view reconstruction**
  - Epipolar geometry
  - Fundamental matrix **F**
  - Essential matrix **E**
  - Computing **F** and **E**

Epipolar Geometry



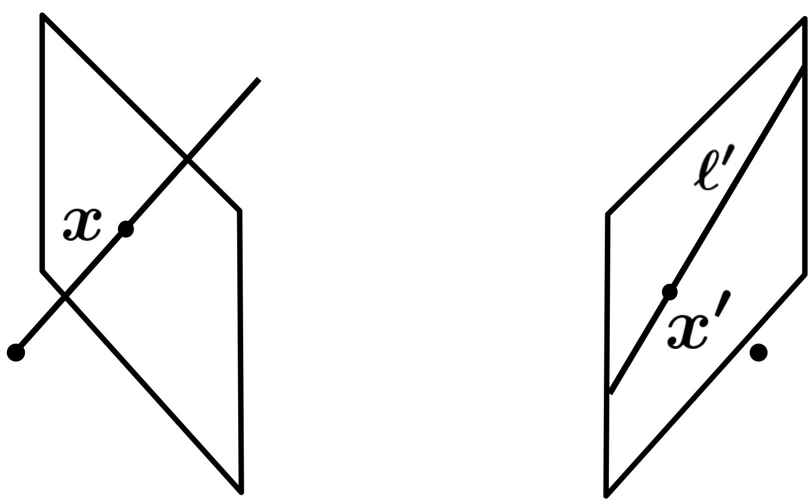


The Fundamental Matrix **F**

- Algebraic representation of epipolar geometry
  - 3x3 Matrix
  - Maps points to epipolar lines

$$\ell' = Fx \qquad \ell = F^T x'$$

  - Epipolar constraint $x'^T F x = 0$

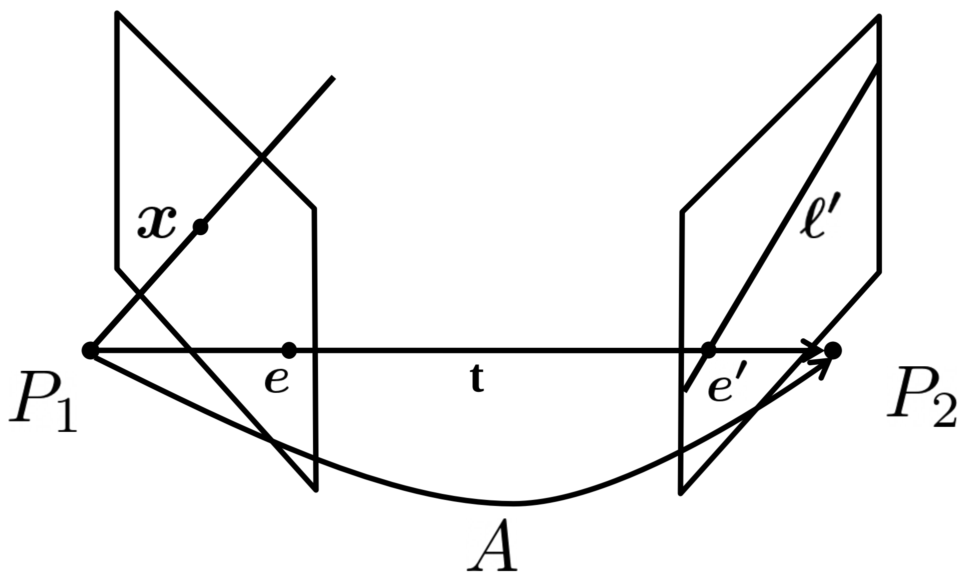


ETH

The Fundamental Matrix **F**

Fundamental matrix encodes relative pose

$$F \quad \Leftrightarrow^{*} \quad \begin{matrix} P_1 = [I \ \mathbf{0}] \\ P_2 = [A \ \boldsymbol{t}] \end{matrix}$$



*up to a projective coordinate change!

Properties of F

F is the unique 3x3 rank 2 matrix that satisfies $x'^T \mathbf{F} x=0$ for all $x \leftrightarrow x'$

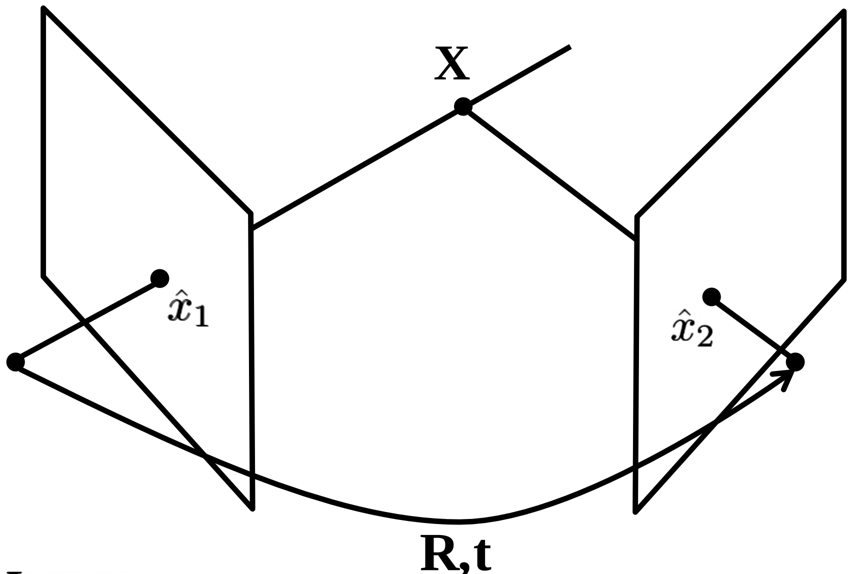
- **Transpose:** if **F** is fundamental matrix for (I,I'), then $\mathbf{F}^T$ is fundamental matrix for (I',I)
- **Epipolar lines:** $l'=\mathbf{F}x$ & $l=\mathbf{F}^T x'$
- **Epipoles:** on all epipolar lines, thus $e'^T \mathbf{F} x=0, \forall x$ $\Rightarrow e'^T \mathbf{F}=0$, similarly $\mathbf{F}e=0$
- **Rank 2:** epipoles in nullspace!
- **F** has 7 degree of freedom (DOF), i.e. 3x3 -1(homogeneous) -1(rank 2)

ETH

The Essential Matrix **E**

- Calibrated case: $P_1 = \cancel{K_1}[I\ 0],\ P_2 = \cancel{K_2}[R\ \boldsymbol{t}]$

$$\hat{x}_i = K_i^{-1} x_i$$



$$\lambda_1 \hat{x}_1 = X$$
$$\lambda_2 \hat{x}_2 = RX + \boldsymbol{t}$$

$$\lambda_2 \hat{x}_2 = \lambda_1 R \hat{x}_1 + \boldsymbol{t}$$

$$(\boldsymbol{t} \times \hat{x}_2)^T \boldsymbol{R} \hat{x}_1 = 0$$
$$([\boldsymbol{t}]_\times \hat{x}_2)^T \boldsymbol{R} \hat{x}_1 = 0$$
$$\hat{x}_2^T [\boldsymbol{t}]_\times^T \boldsymbol{R} \hat{x}_1 = 0$$

$$[\boldsymbol{t}]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$$

$$\boldsymbol{t} \times \hat{x}_2 = [\boldsymbol{t}]_\times \hat{x}_2$$

Properties of **E**

E is an essential matrix iff two of its singular values are equal, third is 0

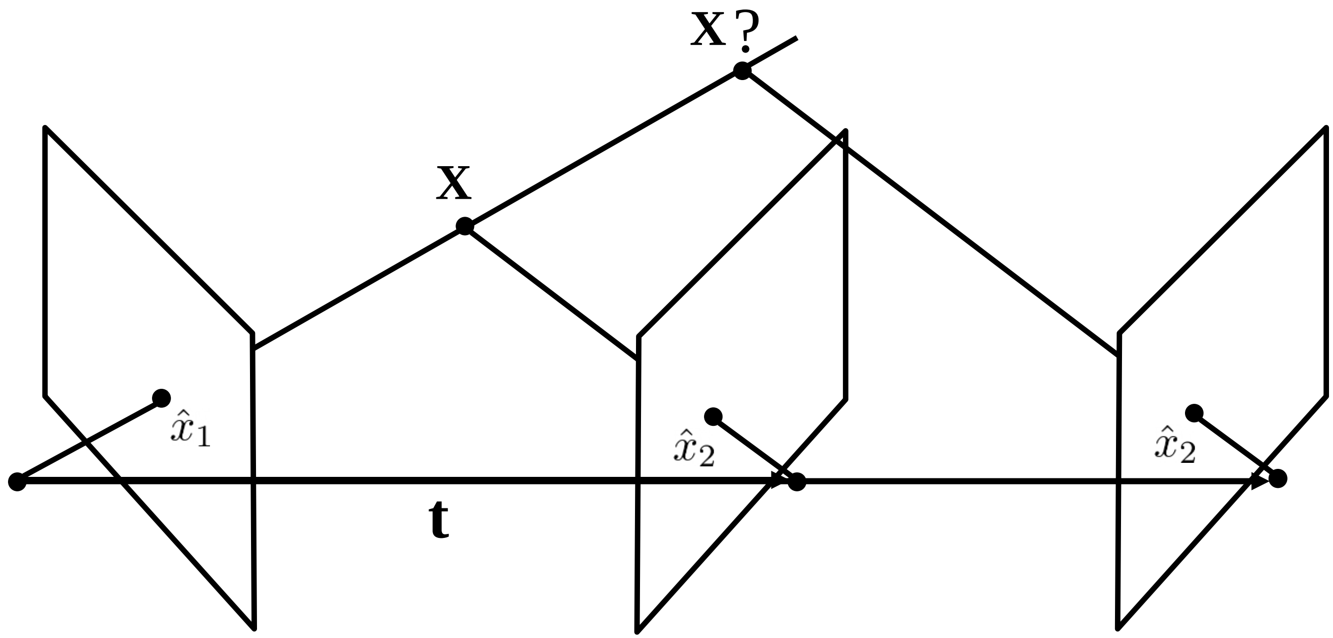
- Relationship to **F**?

$$x_2^T F x_1 = 0 \qquad \hat{x}_2^T E \hat{x}_1 = 0 \qquad \hat{x}_i = K_i^{-1} x_i$$

$$x_2^T \underbrace{K_2^{-T} E K_1^{-1}}_{F} x_1 = 0$$

- Inherits **F**'s properties (see previous slide)

Degree of Freedom of $E = [\mathbf{t}]_\times R$



ETH

Computation of **F** & **E**

- Linear (8-point) (**F** & **E**)
- Minimal (7-point) (**F** & **E**)
- Calibrated (5-point) (only **E**)

Linear Solution (8-point)

- Basic epipolar equation: $\mathrm{x'}^{\mathrm{T}}\mathrm{Fx} = 0$
- Expand:

$$x'xf_{11} + x'yf_{12} + x'f_{13} + y'xf_{21} + y'yf_{22} + y'f_{23} + xf_{31} + yf_{32} + f_{33} = 0$$

- Separate known and unknown variables:

$$[x'x, x'y, x', y'x, y'y, y', x, y, 1][f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33}]^{\mathrm{T}} = 0$$

(data) (unknowns)

- Write as linear equation:

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathrm{f} = 0$$

- 8 unknowns (up to scale): Use 8 points

Normalized 8-point Algorithm

$$\begin{bmatrix} x_1x_1' & y_1x_1' & x_1' & x_1y_1' & y_1y_1' & y_1' & x_1 & y_1 & 1 \\ x_2x_2' & y_2x_2' & x_2' & x_2y_2' & y_2y_2' & y_2' & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_nx_n' & y_nx_n' & x_n' & x_ny_n' & y_ny_n' & y_n' & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0$$

~10000 ~10000 ~100 ~10000 ~10000 ~100 ~100 ~100 1

! Orders of magnitude difference between column of data matrix $\rightarrow$ least-squares yields poor results

- Normalize point coordinates prior to computing **F**
- Same as for the normalized DLT algorithm for homography estimation (see lecture 2)

ETH

The Singularity Constraint

The Singularity Constraint

$$e'^{T} F = 0 \qquad Fe = 0 \qquad \det F = 0 \qquad \text{rank } F = 2$$

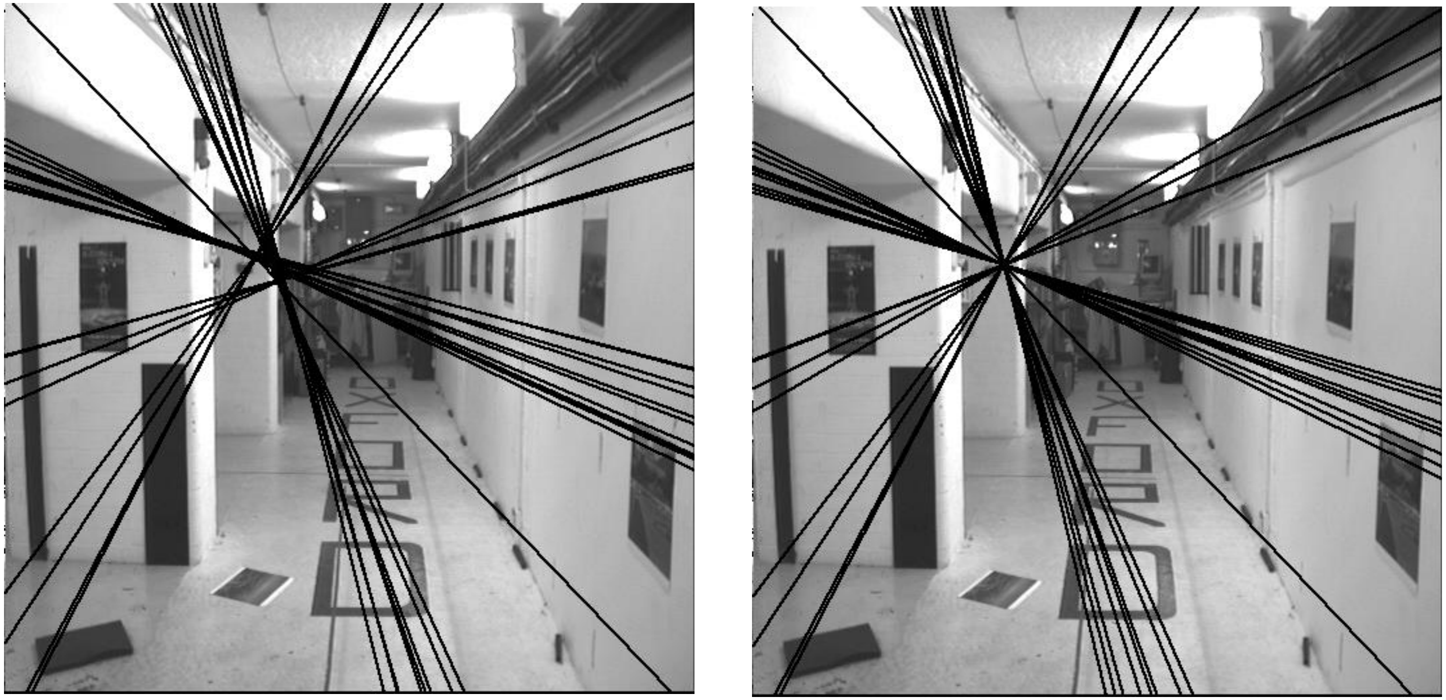
- SVD from linearly computed F matrix (rank 3):

$$F = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T + U_3 \sigma_3 V_3^T$$

- Compute closest rank-2 approximation: $\min \|F - F'\|_F$

$$F' = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & 0 \end{bmatrix} V^T = U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T$$

The Singularity Constraint



Minimal Case: 7 Point Correspondences

- Setup linear system from 7 correspondences:

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_7 x_7 & x'_7 y_7 & x'_7 & y'_7 x_7 & y'_7 y_7 & y'_7 & x_7 & y_7 & 1 \end{bmatrix} \mathrm{f} = 0$$

- Resulting solution has 2D solution space

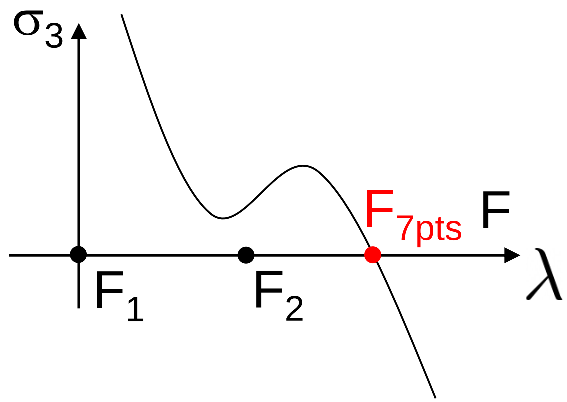
$$\mathrm{A} = \mathrm{U}_{7\mathrm{x}7} \mathrm{diag}(\sigma_1, \ldots, \sigma_7, 0, 0) \mathrm{V}_{9\mathrm{x}9}{}^{\mathrm{T}} \quad \Rightarrow \quad \mathrm{A}[\mathrm{V}_8 \mathrm{V}_9] = 0_{9\mathrm{x}2}$$

- F is linear combination of $\mathbf{V}_8$ and $\mathbf{V}_9$:

$$\mathrm{x}_i{}^{\mathrm{T}} (\mathrm{F}_1 + \lambda \mathrm{F}_2) \mathrm{x}_i = 0, \forall i = 1 \ldots 7$$

  - … but $F_1 + \lambda F_2$ not automatically rank 2

Minimal Case: 7 Point Correspondences



- Enforce rank-2 constraint from determinant:

$$\det(\mathrm{F}_1 + \lambda \mathrm{F}_2) = a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0$$

  - Cubic equation in λ
  - Either 1 or 3 solutions

Calibrated Case: 5-point Solver

D. Nister, An Efficient Solution to the Five-Point Relative Pose Problem, CVPR 2003

- Linear equations from 5 points

$$\begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_{11} & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ x'_2x_2 & x'_2y_2 & x'_{21} & y'_2x_2 & y'_2y_2 & y'_2 & x_2 & y_2 & 1 \\ x'_3x_3 & x'_3y_3 & x'_{31} & y'_3x_3 & y'_3y_3 & y'_3 & x_3 & y_3 & 1 \\ x'_4x_4 & x'_4y_4 & x'_{41} & y'_4x_4 & y'_4y_4 & y'_4 & x_4 & y_4 & 1 \\ x'_5x_5 & x'_5y_5 & x'_{51} & y'_5x_5 & y'_5y_5 & y'_5 & x_5 & y_5 & 1 \end{bmatrix} \begin{bmatrix} E_{11} \\ E_{12} \\ E_{13} \\ E_{21} \\ E_{22} \\ E_{23} \\ E_{31} \\ E_{32} \\ E_{33} \end{bmatrix} = 0$$

- 4D linear solution space:

  E = xX + yY + zZ + wW   scale does not matter, choose w = 1

- Insert into non-linear constraints

$$\det E = 0$$

$$2\mathbf{E}\mathbf{E}^T\mathbf{E} - tr(\mathbf{E}\mathbf{E}^T)\mathbf{E} = 0.$$

} 10 cubic polynomials

(assumes normalized coordinates)

ETH

Calibrated Case: 5-point Solver

D. Nister, An Efficient Solution to the Five-Point Relative Pose Problem, CVPR 2003

- Perform Gauss-Jordan elimination on polynomials

[n] represents polynomial of degree n in z

	A	x^3	y^3	x^2y	xy^2	x^2z	x^2	y^2z	y^2	xyz	xy	x	y	1
	$\langle a\rangle$	1	.	.	.	.	.	.	.	.	.	[2]	[2]	[3]
	$\langle b\rangle$		1	.	.	.	.	.	.	.	.	[2]	[2]	[3]
	$\langle c\rangle$			1	.	.	.	.	.	.	.	[2]	[2]	[3]
	$\langle d\rangle$				1	.	.	.	.	.	.	[2]	[2]	[3]
$\langle k\rangle$	$\langle e\rangle$					1						[2]	[2]	[3]
-z	$\langle f\rangle$						1					[2]	[2]	[3]
$\langle l\rangle$	$\langle g\rangle$							1				[2]	[2]	[3]
-z	$\langle h\rangle$								1			[2]	[2]	[3]
$\langle m\rangle$	$\langle i\rangle$									1		[2]	[2]	[3]
-z	$\langle j\rangle$										1	[2]	[2]	[3]

$$\langle k\rangle \equiv \langle e\rangle - z\langle f\rangle$$

$$\langle l\rangle \equiv \langle g\rangle - z\langle h\rangle$$

$$\langle m\rangle \equiv \langle i\rangle - z\langle j\rangle$$

B	x	y	1
$\langle k\rangle$	[3]	[3]	[4]
$\langle l\rangle$	[3]	[3]	[4]
$\langle m\rangle$	[3]	[3]	[4]

$$\langle n\rangle \equiv det(B)$$

Automatic Computation of **F**

Step 1. Extract features

Step 2. Compute a set of potential matches

Step 3. **Robust estimation of F via RANSAC**

Step 4. Compute F based on all inliers

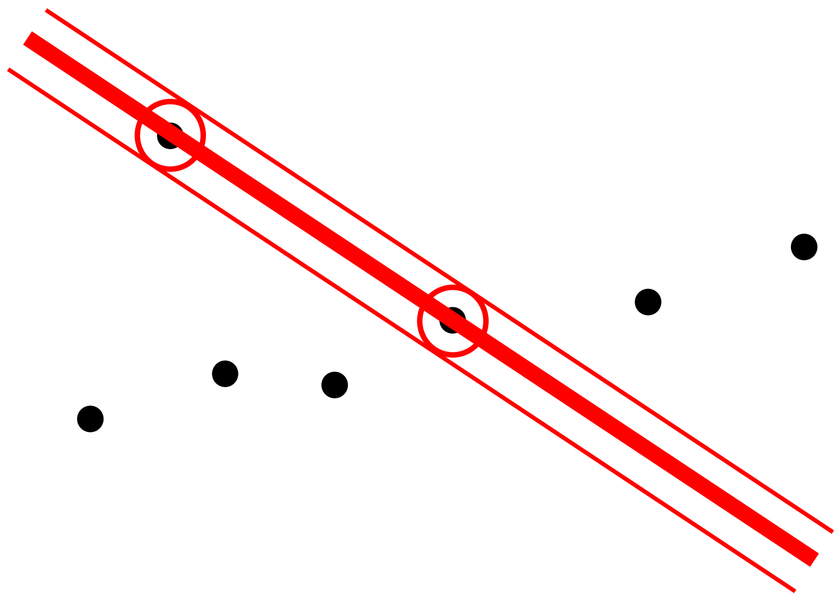
Step 5. **Look for additional matches**

Step 6. Refine F based on all correct matches

ETH

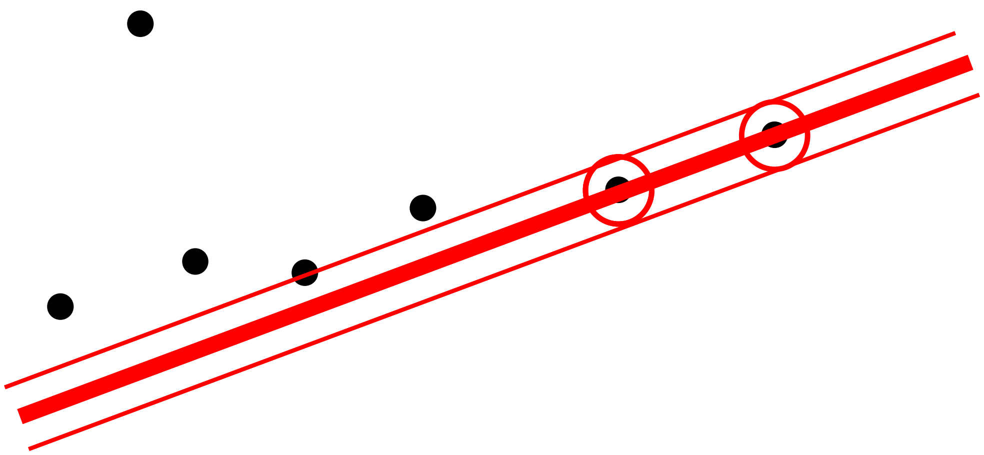
RANDom SAmple Consensus (RANSAC)

M. A.. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, CACM 1981



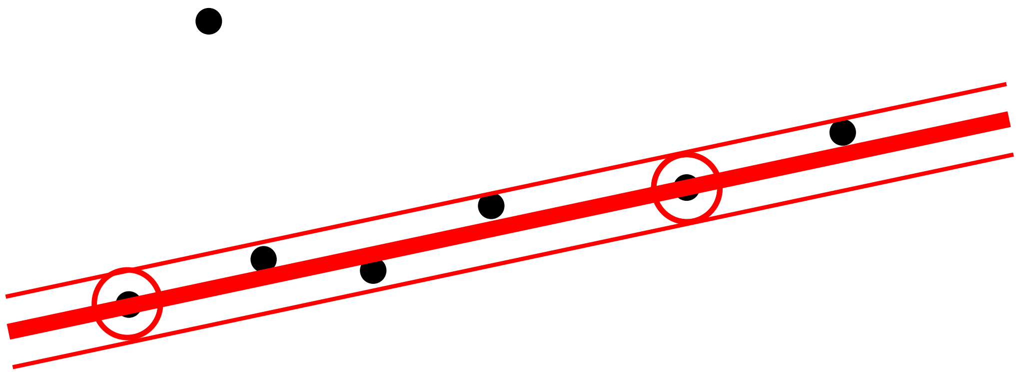
RANdom SAmple Consensus (RANSAC)

M. A.. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, CACM 1981



RANdom SAmple Consensus (RANSAC)

M. A.. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, CACM 1981



RANdom SAmple Consensus (RANSAC)

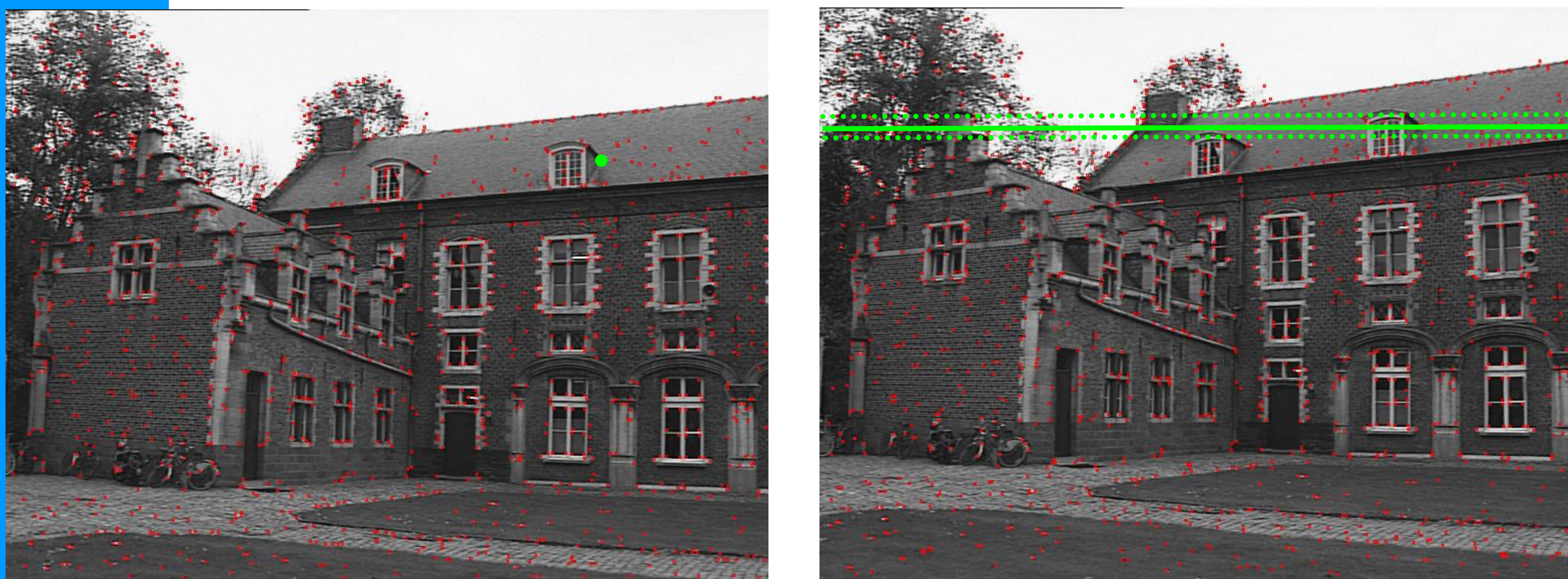
M. A.. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, CACM 1981

- Problem: Estimate **F** in presence of wrong matches
- RANSAC algorithm:
  - Repeat:
    - Randomly select minimal sample (5 or 7 points)
    - Compute hypothesis for **F** from minimal sample
    - Verify hypothesis: Count inliers
    - Update best solution found so far
  - Until prob. of not having sampled all-inlier set< η

% inliers	90%	80%	70%	60%	50%	20%
#samples (5)	5	12	25	57	145	14k
#samples (7)	7	20	54	162	587	359k

η=0.01%

Finding More Matches



- Restricted search around epipolar line
  (e.g. ±1.5 pixels)

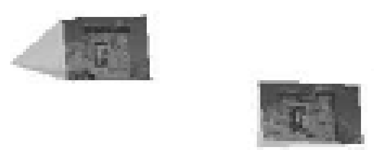
Summary: Epipolar Geometry

http://danielwedge.com/fmatrix/

ETH

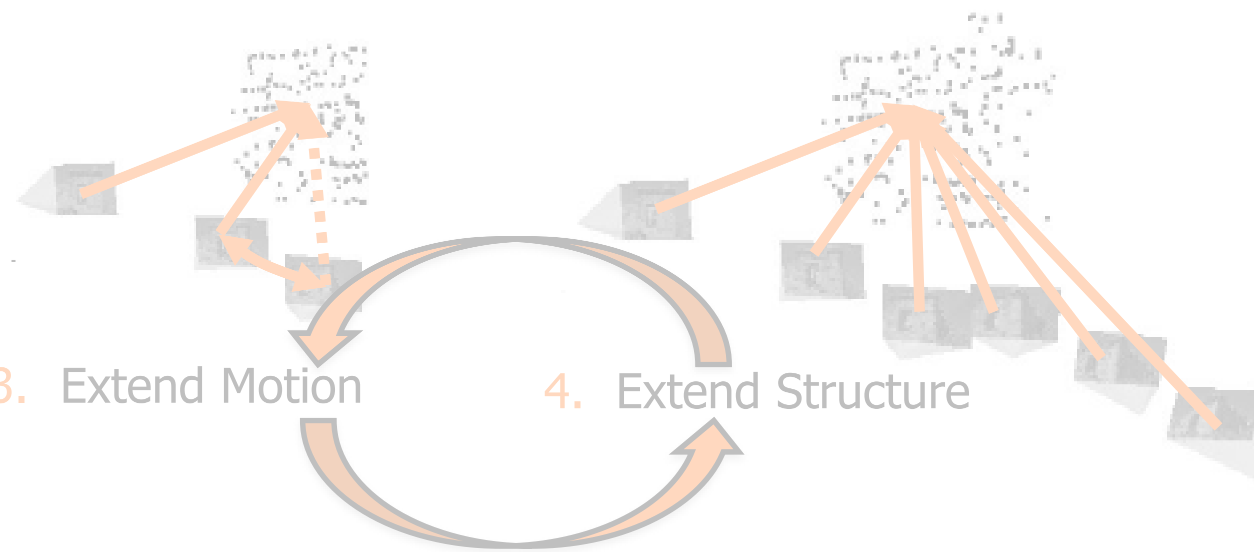
Sequential / Incremental SfM

- Initialize motion from **F** or **E**
- Triangulate structure from motion



1. Initialize Motion

2. **Initialize Structure**



Initial Motion and Structure Estimation (Calibrated Case)

- Recap Essential matrix: $\mathbf{E}=[\mathbf{t}]_x\mathbf{R}$
- Motion for two cameras: $[\mathbf{I}|\mathbf{0}]$, $[\mathbf{R}|\mathbf{t}]$
- Essential Matrix decomposition: $\mathbf{E}=\mathbf{U\Sigma V}^T$

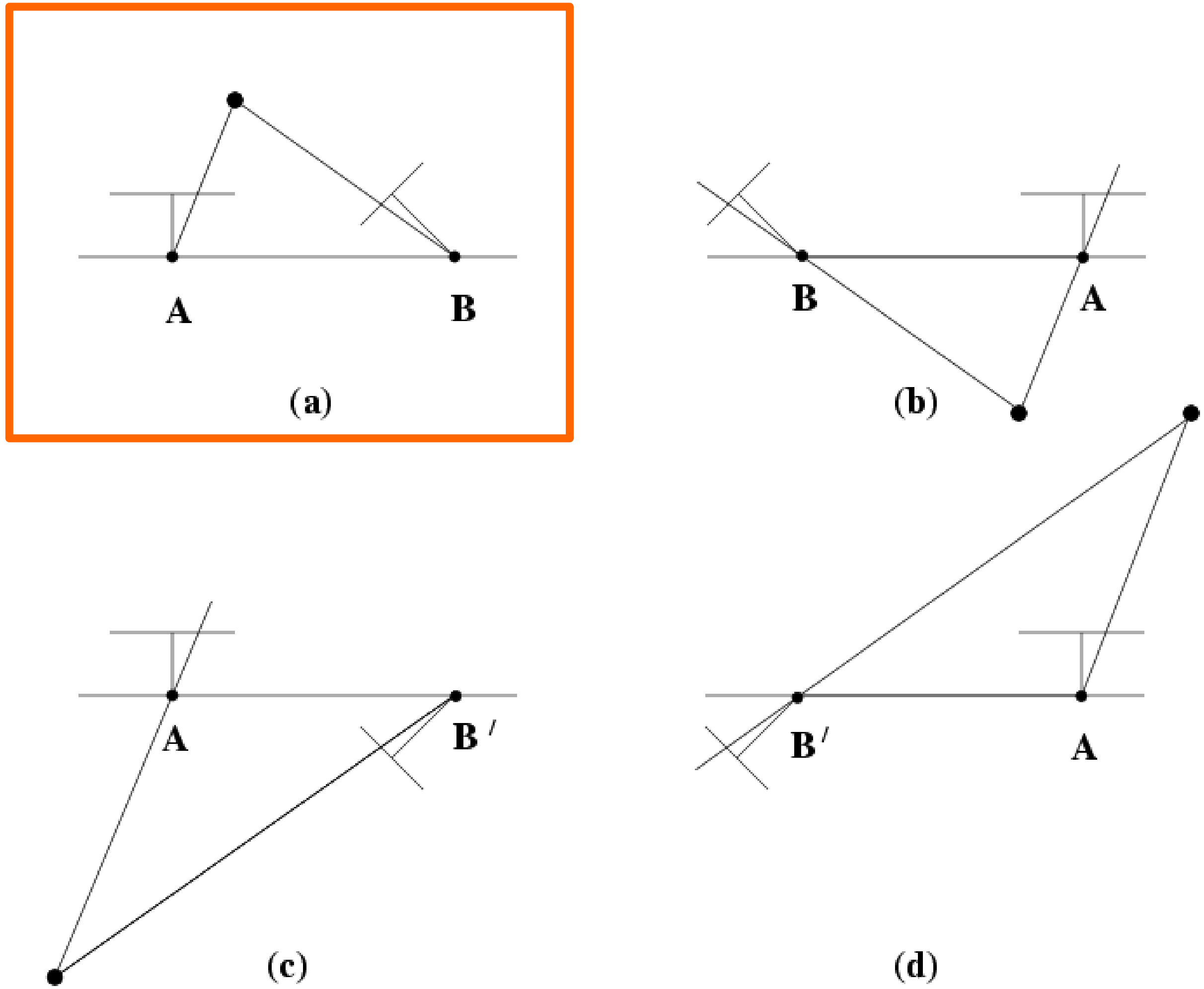
$$\mathbf{\Sigma} = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{W} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{W}^{-1} = \mathbf{W}^T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Recover **R** and **t** as
  - $\mathbf{t}=\mathbf{u}_3$ or $\mathbf{t}=-\mathbf{u}_3$
  - $\mathbf{R}=\mathbf{UWV}^T$ or $\mathbf{R}=\mathbf{UW}^T\mathbf{V}^T$
  - Four solutions, but only one meaningful

(see Hartley and Zisserman, Sec.9.6)

ETH

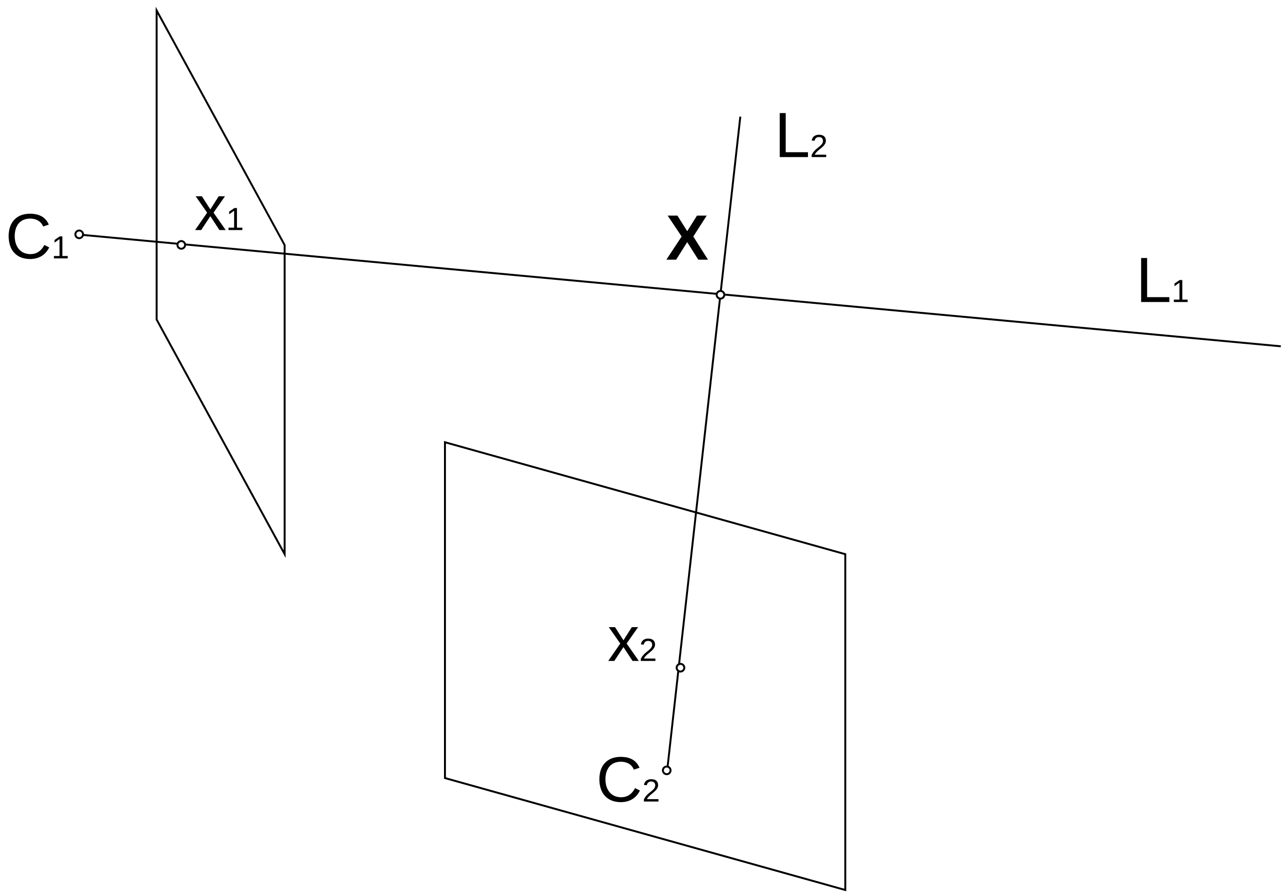
Using the Cheirality Constraint



(see Hartley and Zisserman, Sec. 9.6)

Triangulation

- Given: Motion, correspondence
- Estimate 3D point via **triangulation**



Triangulation

- Backprojection

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

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathtt{P}_1 \\ \mathtt{P}_2 \\ \mathtt{P}_3 \end{bmatrix} \mathtt{X}$$

$$\begin{aligned} \mathtt{P}_3\mathtt{X}x &= \mathtt{P}_1\mathtt{X} \\ \mathtt{P}_3\mathtt{X}y &= \mathtt{P}_2\mathtt{X} \end{aligned}$$

$$\begin{bmatrix} \mathtt{P}_3 x - \mathtt{P}_1 \\ \mathtt{P}_3 y - \mathtt{P}_2 \end{bmatrix} \mathtt{X} = 0$$

- Triangulation

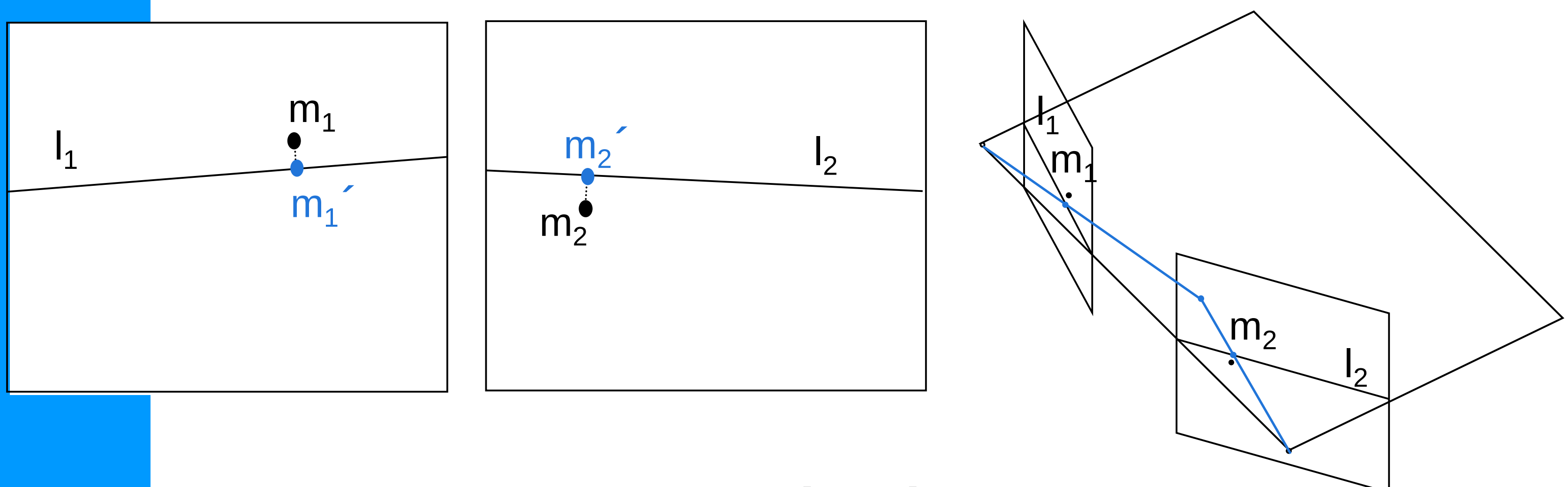
$$\begin{bmatrix} \mathtt{P}_3 x - \mathtt{P}_1 \\ \mathtt{P}_3 y - \mathtt{P}_2 \\ \mathtt{P}'_3 x' - \mathtt{P}'_1 \\ \mathtt{P}'_3 y' - \mathtt{P}'_2 \end{bmatrix} \mathtt{X} = 0$$

- Maximum Likelihood Triangulation (geometric error)

$$\arg\min_{\mathtt{X}} \sum_i \left(\mathtt{x}_i - \lambda^{-1}\mathbf{P}_i\mathtt{X}\right)^2$$

Optimal 3D Point in Epipolar Plane

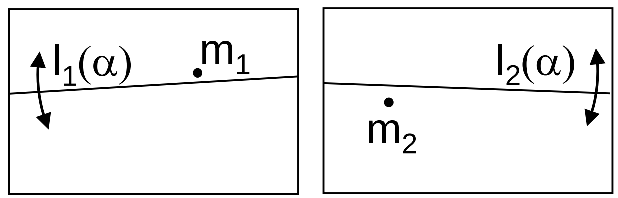
- Given an epipolar plane, find best 3D point for (m_1,m_2)



- Select closest points $(m_1´,m_2´)$ on epipolar lines
- Obtain 3D point through exact triangulation
- Guarantees minimal reprojection error (given this epipolar plane)

Optimal Two-View Triangulation

- Non-iterative method: (Hartley and Sturm, CVIU´97)
  - Determine optimal epipolar plane for reconstruction



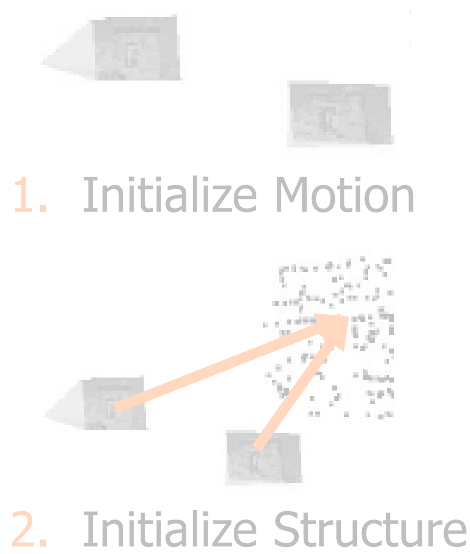
1DOF

$$D\left(\mathbf{m}_1,\mathbf{l}_1(\alpha)\right)^2 + D\left(\mathbf{m}_2,\mathbf{l}_2(\alpha)\right)^2$$ (polynomial of degree 6)

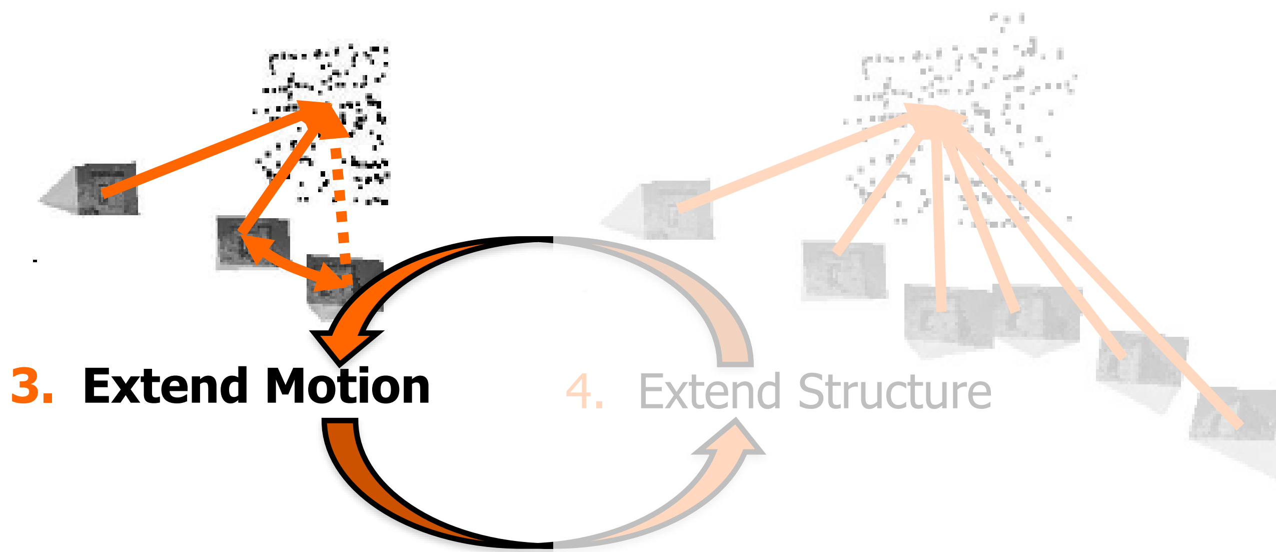
  - Reconstruct optimal point from selected epipolar plane
- Note: Only works for two views

ETH

Sequential / Incremental SfM

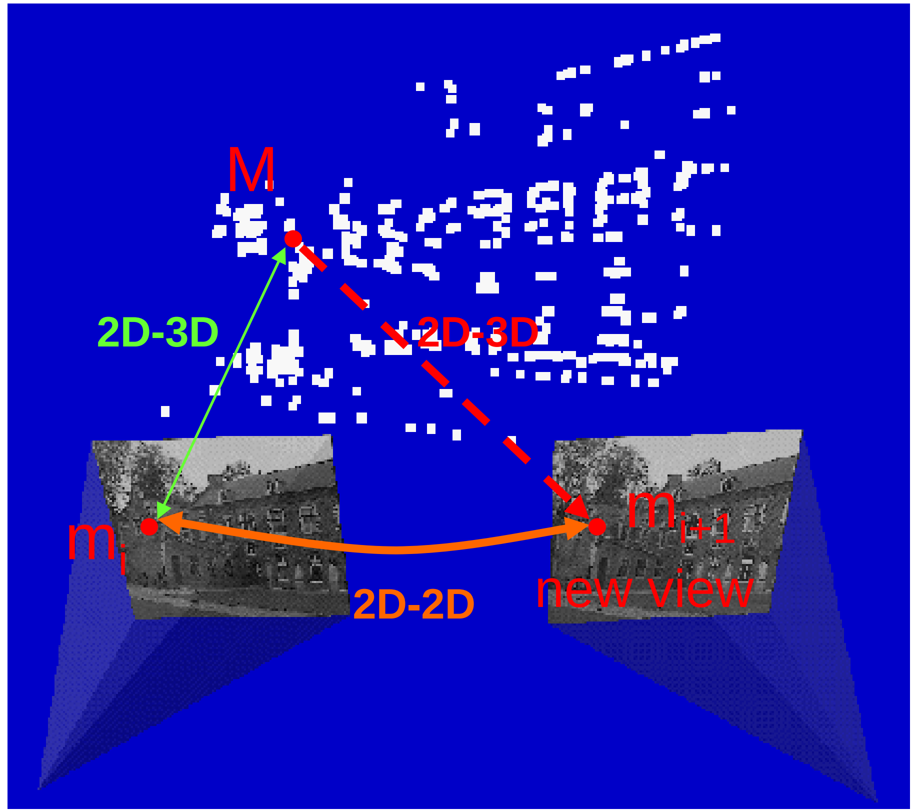


- Find camera with matches to previous images
- Matches define 2D-3D correspondences
- Estimate camera pose wrt. 3D structure



ETH

Pose Estimation from 2D-3D Matches



- Compute $\mathbf{P}_{i+1}$ using robust approach (6-point RANSAC)
- Extend and refine reconstruction

Compute P with 6-point RANSAC

- Generate hypothesis using 6 points

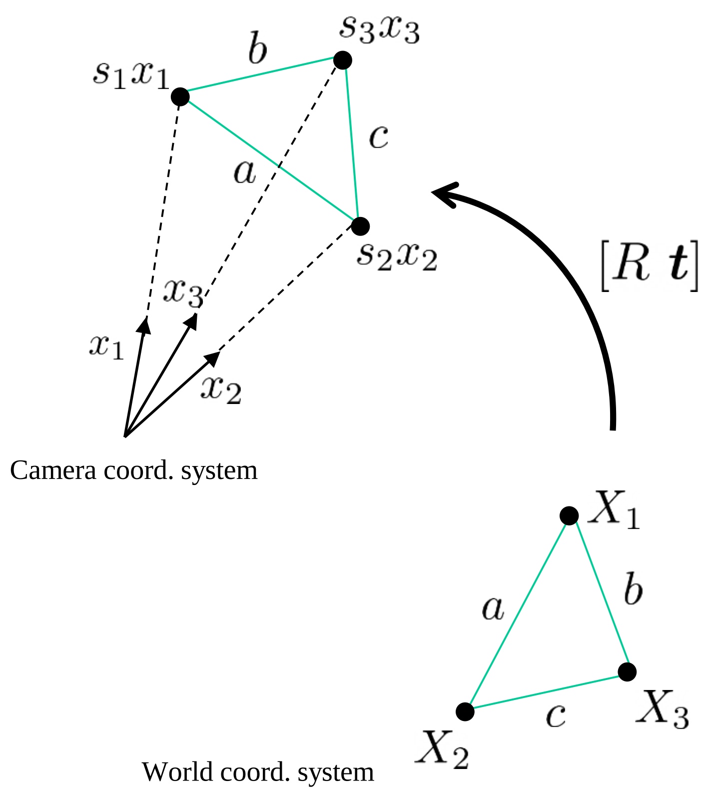
$$\begin{bmatrix} \mathbf{0}^\top & -w_i \mathbf{X}_i{}^\top & y_i \mathbf{X}_i{}^\top \\ w_i \mathbf{X}_i{}^\top & \mathbf{0}^\top & -x_i \mathbf{X}_i{}^\top \end{bmatrix} \begin{pmatrix} \mathbf{P}^1 \\ \mathbf{P}^2 \\ \mathbf{P}^3 \end{pmatrix} = \mathbf{0}$$

(two equations per point)

- Planar scenes are degenerate!

(similar DLT algorithm as see in 2nd lecture for homographies)

3-Point-Perspective Pose – P3P (Calibrated Case)



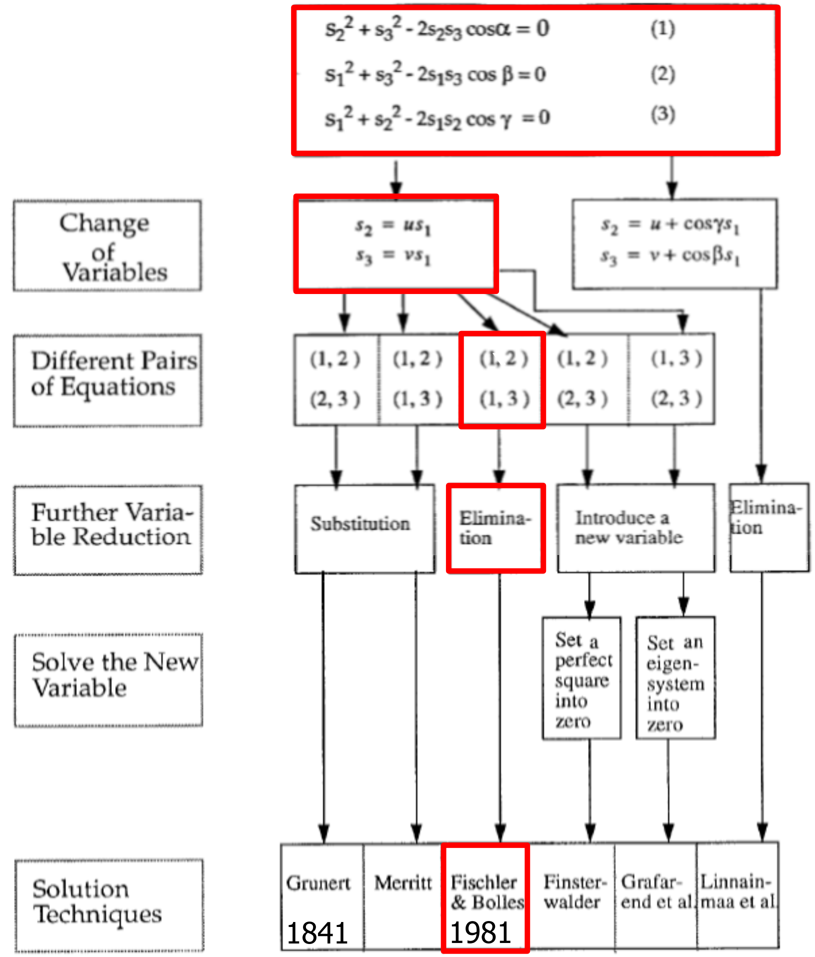


Fig. 2. Shows the differences of a algebraic derivations among six solution techniques.

All techniques yield 4th order polynomial

R. Haralick, D. Lee, K. Ottenburg, M. Nolle. Review and analysis of solutions of the three point perspective pose estimation problem, CVPR 1991

ETH

Incremental SfM

- **Initialize:**
  - Compute pairwise epipolar geometry
  - Find pair to initialize structure and motion
- **Repeat:**
  - For each additional view
    - Determine pose from structure
    - Extend structure
    - Refine structure and motion (bundle adjustment, see lecture 7)

Global SfM

- **Initialize:**
  - Compute pairwise epipolar geometry
- **Compute:**
  - Estimate all orientations
  - Estimate all positions
  - Triangulate structure
  - Refine structure and motion (bundle adjustment)
- **Pros**: More efficient, more accurate
- **Con**: Less robust

SfM Software

- Colmap (Johannes Schönberger)
  - Incremental SfM, very efficient, nice GUI, open source
- VisualSFM (Changchang Wu)
  - Incremental SfM, very efficient, GUI, binaries
- Bundler (Noah Snavely)
  - Incremental SfM, open source
- OpenMVG (Pierre Moulon)
  - Incremental and Global SfM, open source
- Theia (Chris Sweeney)
  - Incremental and Global SfM, very efficient, open source

Summary

- Estimate motion between two images
  - Epipolar geometry
- Estimate structure from motion
  - Triangulation
- Estimate camera pose from structure
  - Absolute camera pose solvers
    - DLT 6-point solver (P6P)
    - 3-point-perspective pose solver (P3P)

Schedule

Feb 19	Introduction
Feb 26	Geometry, Camera Model, Calibration
Mar 4	Guest lecture + Features, Tracking / Matching
Mar 11	**Project Proposals by Students**
Mar 18	3DV conference
Mar 25	**Structure from Motion (SfM) + papers**
Apr 1	Easter break
Apr 8	Dense Correspondence (stereo / optical flow) + papers
Apr 15	Bundle Adjustment & SLAM + papers
Apr 22	**Student Midterm Presentations**
Apr 29	Multi-View Stereo & Volumetric Modeling + papers
May 6	3D Modeling with Depth Sensors + papers
May 13	Guest lecture + papers
May 20	Holiday

ETH

Next week:
Dense Correspondence / Stereo

Now:
Paper presentations!