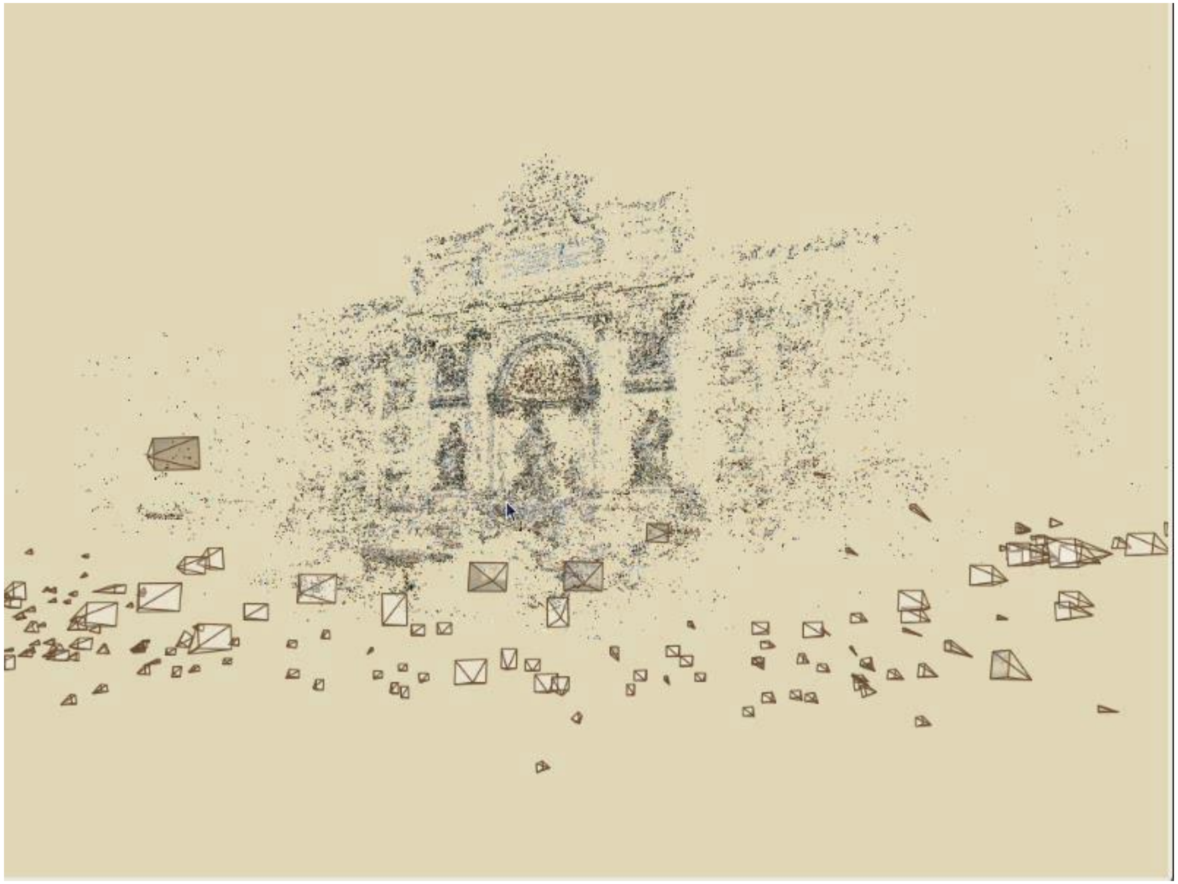
Class 6



Last Class

- Camera calibration
- General case triangulation

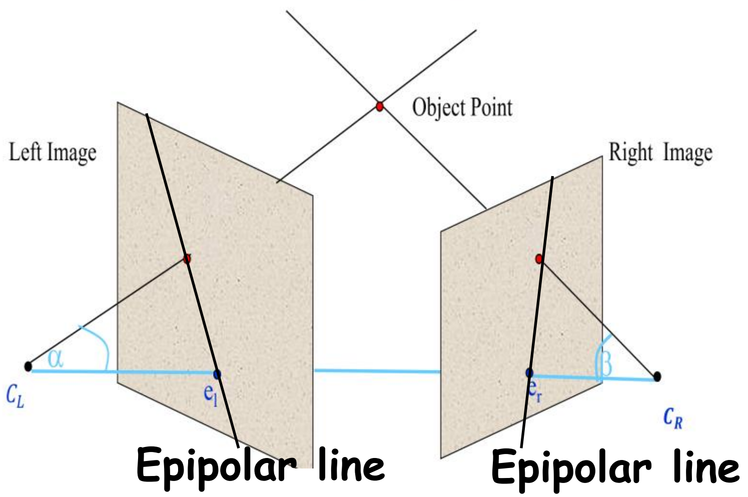
$\tilde{p} = M\,\tilde{P}$

- Epipolar geometry: $\tilde{p}_R^T F \tilde{p}_L = 0$

Epipolar Geometry

A pair of corresponding points must lay on corresponding epipolar lines

- $\tilde{p}_R^T F \tilde{p}_L = 0$
- $F = [\mathrm{e_r}]_\times M_R M_L^+$
- $\tilde{\ell}_R = F\tilde{p}_L$ & $\tilde{\ell}_L = \tilde{p}_R^T F$
- The epipoles:
  - $\tilde{e}_R^T F = 0$ & $F\tilde{e}_L = 0$



A Special Case: Pure Translation

$$F = [\mathrm{e_r}]_\times M_R M_L^+$$

Assume:

- The world coordinate system is the left camera's system
- Both cameras have the same intrinsic projection matrix

$$M_{int} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

- The right camera is only translated in the x direction

Then:

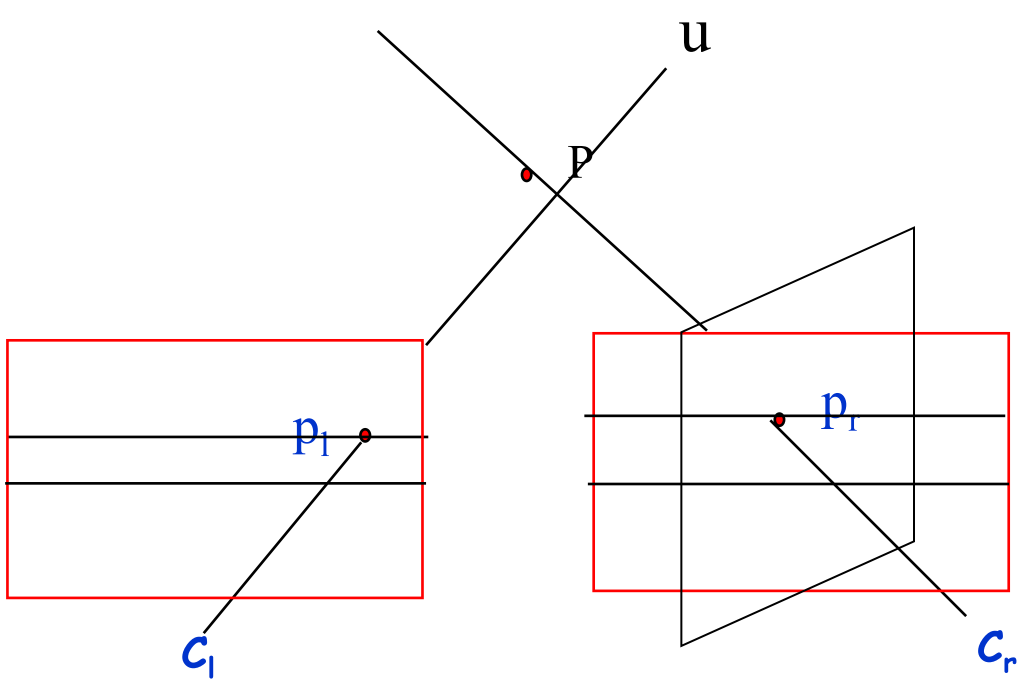
$$F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix}$$

?

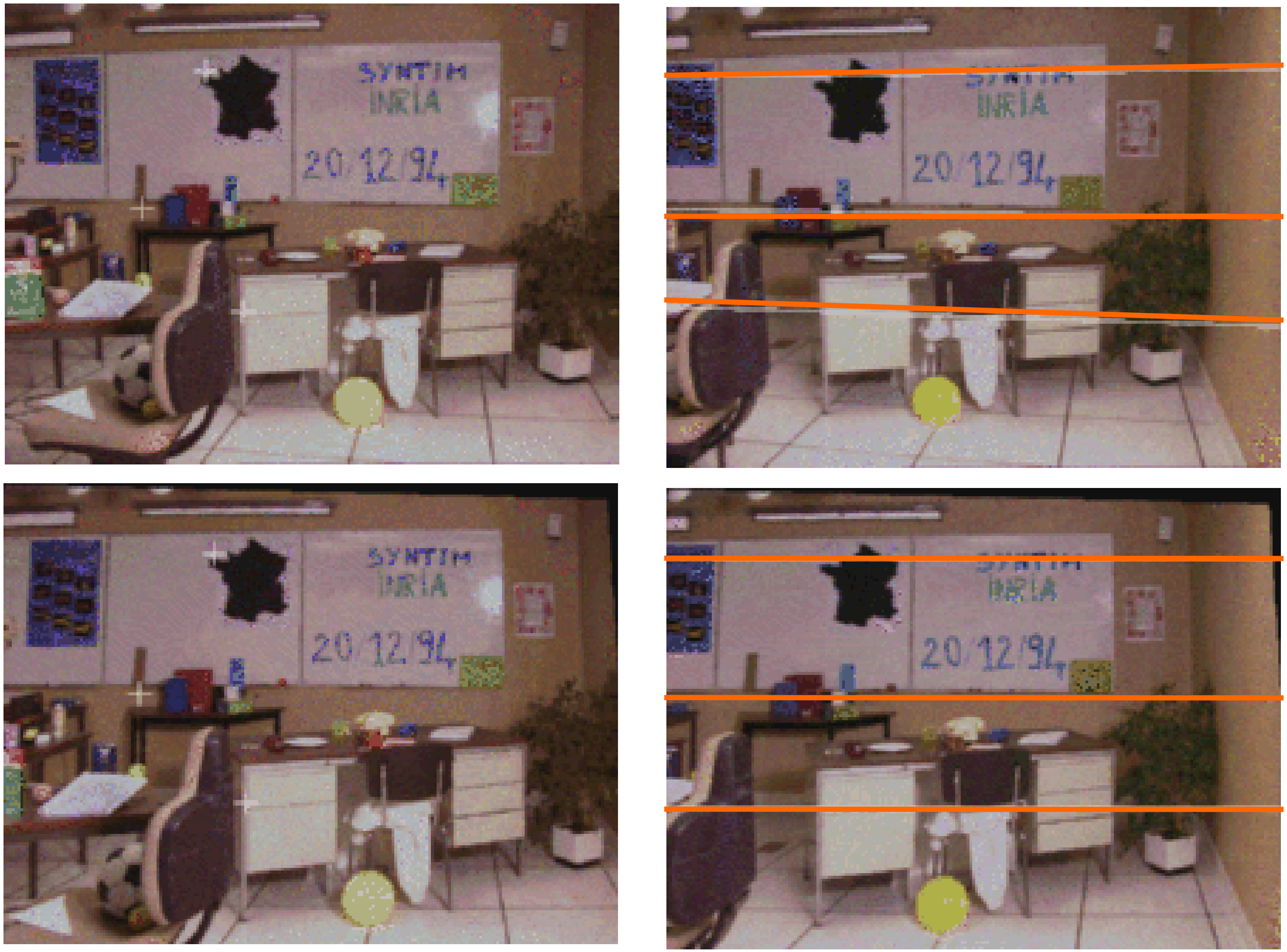
$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -t_x \\ 0 & t_x & 1 \end{pmatrix}$$

Rectification

- Simplify the use of epipolar lines



Rectification



Taken from: SYNTIM-INRIA.

Rectification

- Given $\tilde{p}_R^T F \tilde{p}_L$
- Let *A* and *B* be *3×3* non singular matrices
- Transform the left image by $\tilde{q}_L = A\tilde{p}_L$
- Transform the right image by $\tilde{q}_R^T = B\,\tilde{p}_R$
- $\tilde{p}_R^T F \tilde{p}_L = (B^{-1}\tilde{q}_R^T)^T F (A^{-1}\tilde{q}_L^T) = \tilde{q}_R^T (B^{-T} F A^{-1}) \tilde{q}_L$
- Choose *A* and *B* such that $F' = B^{-T} F A^{-1}$ is a pure translation:

$$F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{pmatrix}$$

Denote: $B^{-T} = (B^{-1})^T$

Computing F

- Given the internal and external parameters of the cameras:

$$F = [\mathrm{e_r}]_\times M_R M_L^+$$

- How many parameters defines F?
- When the cameras' parameters are unknown: use 8 (or more) corresponding points

Computing F

- Assume $n \geq 8$ corresponding points are given
- Use these points to solve the 8 unknowns of F based on: $\tilde{p}_R^T F \tilde{p}_L = 0$
- Avoid degenerate configurations

Eight Points Algorithm

- Each pair of corresponding points defines one linear equation in 8 unknowns

$$(u,v,1)\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \Leftrightarrow (uu', uv', u, vu', vv', v, u', v', 1)\begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0.$$

הבינתחומי הרצליה
IDC HERZLIYA

8 pairs of corresponding points define eight linear equations in 8 unknowns

$$\begin{pmatrix} u_1u_1' & u_1v_1' & u_1 & v_1u_1' & v_1v_1' & v_1 & u_1' & v_1' \\ u_2u_2' & u_2v_2' & u_2 & v_2u_2' & v_2v_2' & v_2 & u_2' & v_2' \\ u_3u_3' & u_3v_3' & u_3 & v_3u_3' & v_3v_3' & v_3 & u_3' & v_3' \\ u_4u_4' & u_4v_4' & u_4 & v_4u_4' & v_4v_4' & v_4 & u_4' & v_4' \\ u_5u_5' & u_5v_5' & u_5 & v_5u_5' & v_5v_5' & v_5 & u_5' & v_5' \\ u_6u_6' & u_6v_6' & u_6 & v_6u_6' & v_6v_6' & v_6 & u_6' & v_6' \\ u_7u_7' & u_7v_7' & u_7 & v_7u_7' & v_7v_7' & v_7 & u_7' & v_7' \\ u_8u_8' & u_8v_8' & u_8 & v_8u_8' & v_8v_8' & v_8 & u_8' & v_8' \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

Solve the linear equations and find F

Improve the 8 points algorithm

- F has rank 2
- Enforce the rank 2
- E.g. Hartley 1995

RANSAC

- RAndom SAmple Consensus
- An iterative method to estimate parameters of a model from a set of observed data which contains outliers, e.g.,
  - A set of points → a line
  - A set of corresponding points → fundamental matrix
  - A set of corresponding points → Homography

RANSAC

- Assumption:
  - The model can be computed from a small set of points
  - The set of points consists of inliers and outliers
- Challenges:
  - Which points to use?
  - When to stop?

video

RANSAC

- Iteratively selecting a random subset of points:
  - Fit a model
  - Test with the entire set: inliers and outliers
  - Support set: sufficient large number of inliers
- Repeat the model computation

RANSAC – Computing F

- Compute corresponding pairs between a pair of images
- Iteratively selecting a random subset of pairs:
  - Fit a model – compute F
  - Test with the entire set: inliers and outliers
  - Support set: sufficient large number of inliers
- Repeat the model computation

Data Sets

- Classic:
  - https://vision.middlebury.edu/stereo/data/
- CMP Extreme View Dataset:
  - http://cmp.felk.cvut.cz/wbs/
- Multi view stereo:
  - http://grail.cs.washington.edu/projects/mvscpc/

Uncalibrated Stereo

- Calibration is necessary to determine absolute 3D positions
- We can determine relative 3D positions (up to a scale factor) without calibration

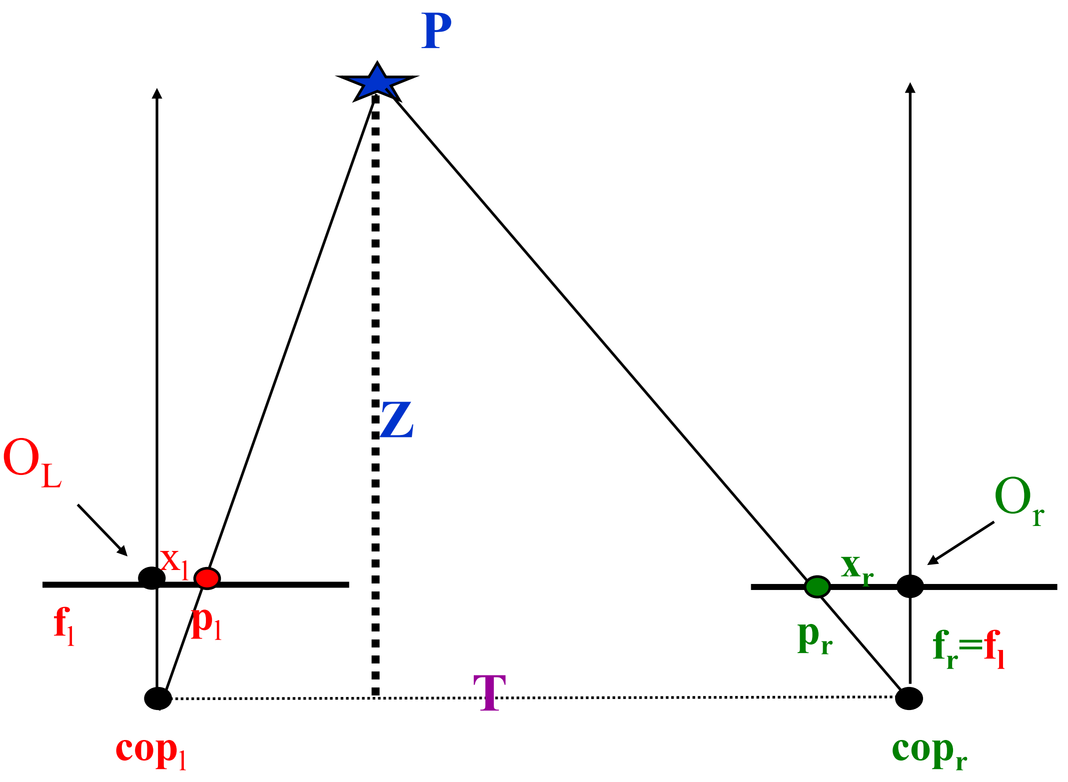
What if we do not know the external parameters?

The disparity:

$$d = x_l - x_r$$

The depth:

$$z = f\,\frac{T}{d}$$



A Priory Knowledge	3D-Reconstruction From 2 Views
Intrinsic and extrinsic parameters	Unambiguous
Intrinsic parameters	Up to unknown scaling factor
No information about parameters	Up to unknown projective transformation

Occlusion Aware

- Seam Carving

Stereo **seam carving** a geometrically consistent approach (Dekel. et al 2013)

Next

- Special cases - Homography
  - planar surfaces, camera rotation

- Other stereo pairs

- More than 2 images:
  - Structure from motion

Questions

- Given a single image:
  - Can we generate another image of the same scene?
  - If so, which?

IDC HERZLIYA הבינתחומי הרצליה

Homography

- A homography transformation:
  Let $p \in I_1, q \in I_2$ be corresponding points
  $\tilde{p} = H\tilde{q}$, where H is a 3×3 matrix
- A pair of perspective images are related by an Homography transformation:
  - Scene is a planar surface
  - The cameras are in the same location (identical up to rotation and intrinsic parameters)

Planar Surface

$$\boxed{aP_x + bP_y + cP_z + d = 0}$$

$$\underset{3\times 4}{M}\begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11}P_x + m_{12}P_y + m_{13}P_z + m_{14} \\ \dots \\ .. \end{pmatrix}$$

w.l.g.:
$P_z = -d$

$$= \begin{pmatrix} m_{11}P_x + m_{12}P_y - m_{13}d + m_{14} \\ \dots \\ .. \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} m_{11} & m_{12} & -m_{13}\,d + m_{14} \\ m_{21} & m_{12} & -m_{23}\,d + m_{24} \\ m_{31} & m_{32} & -m_{33}\,d + m_{34} \end{pmatrix}}_{\substack{A \\ 3\times 3}} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix}$$

IDC HERZLIYA הבינתחומי הרצליה

Planar Surface

$$aP_x + bP_y + cP_z + d = 0$$

$$\text{w.l.g} : P_z = -d$$

- $\tilde{p} = M_1 \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} = \underset{3\times 3}{A} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} = A^{-1}\tilde{p}$
- $\tilde{q} = M_2 \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} = \underset{3\times 3}{B} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \rightarrow \tilde{q} = BA^{-1}\tilde{p} = \underset{3\times 3}{H}\tilde{p}$

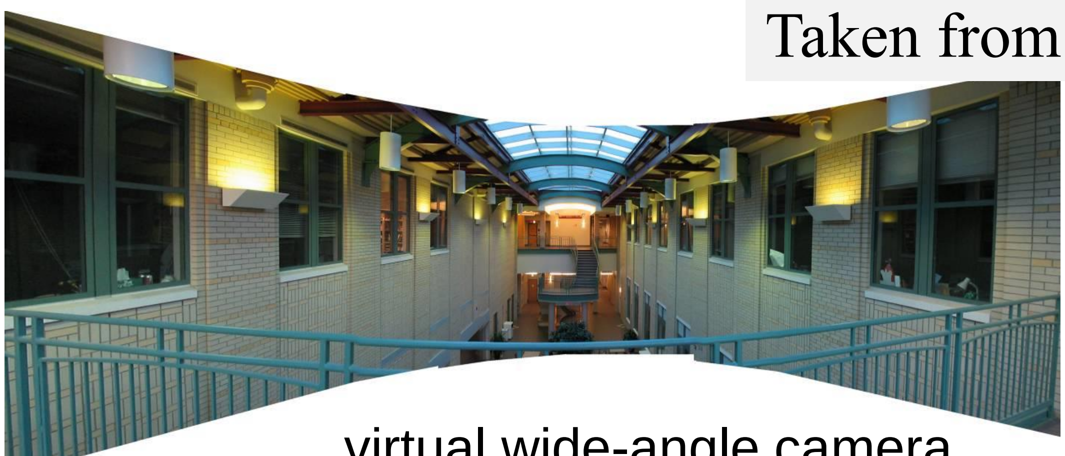
What id the rank of H?

Homography

- Computing using
  - Projection matrices (known plane)
  - 4 corresponding points: $\tilde{q} = H\tilde{p}$
  - RANSAC
- Applications:
  - Mosaics: stitching images together
  - 3D structure

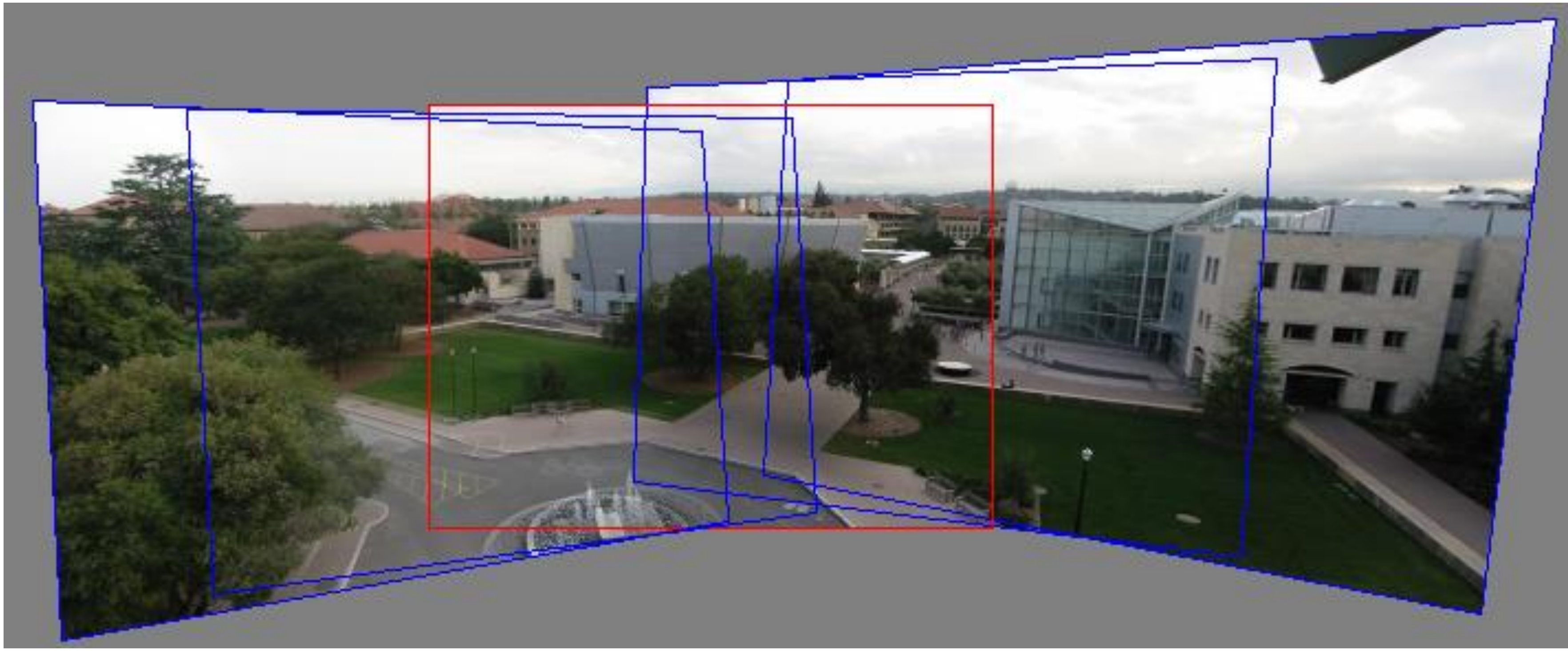
Homography: Pure Rotation

Taken from Jeffrey Martin



virtual wide-angle camera

28

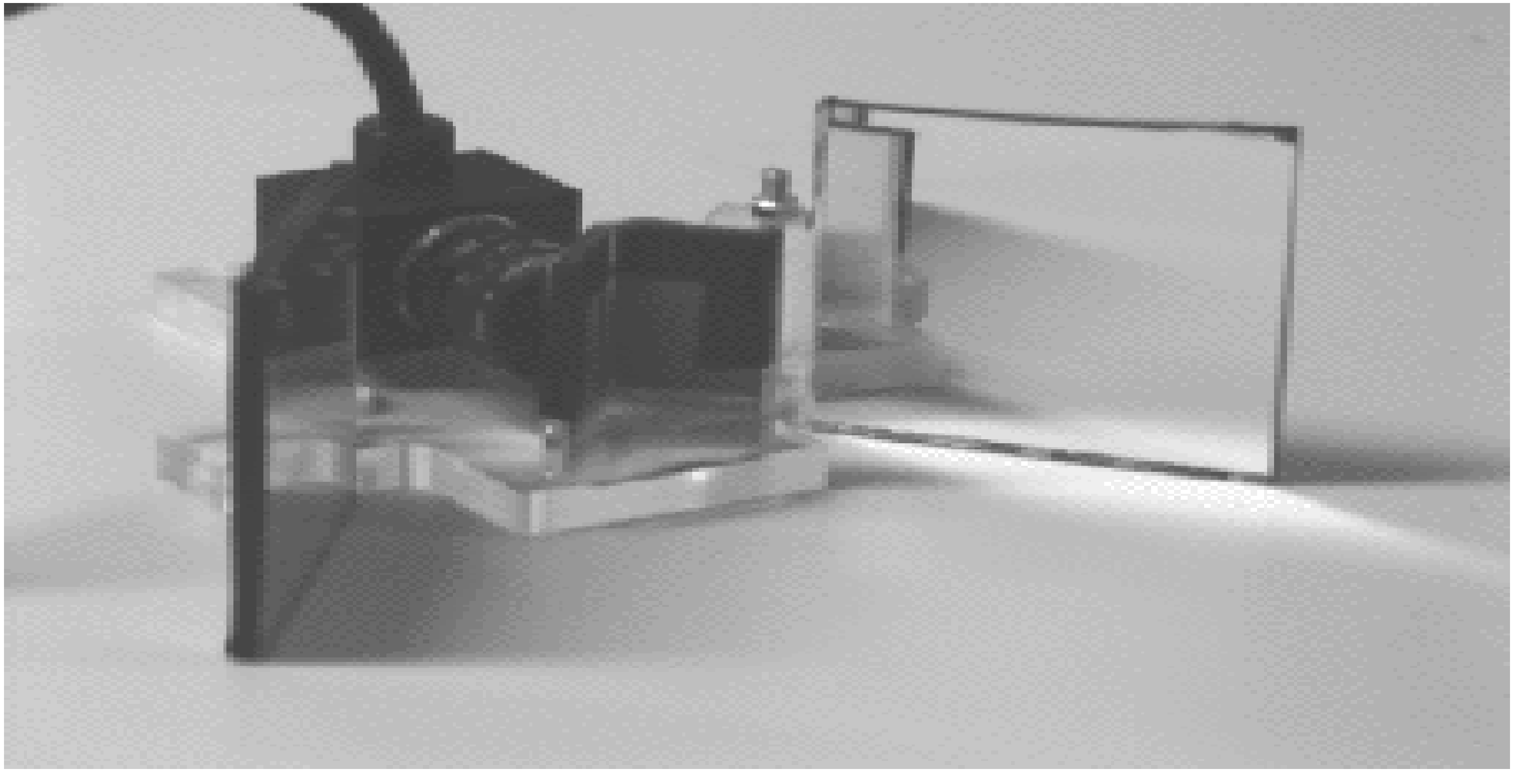


1. Pick one image (red)
2. Warp the other images towards it (usually, one by one)
3. Blend

Taken from Jeffrey Martin

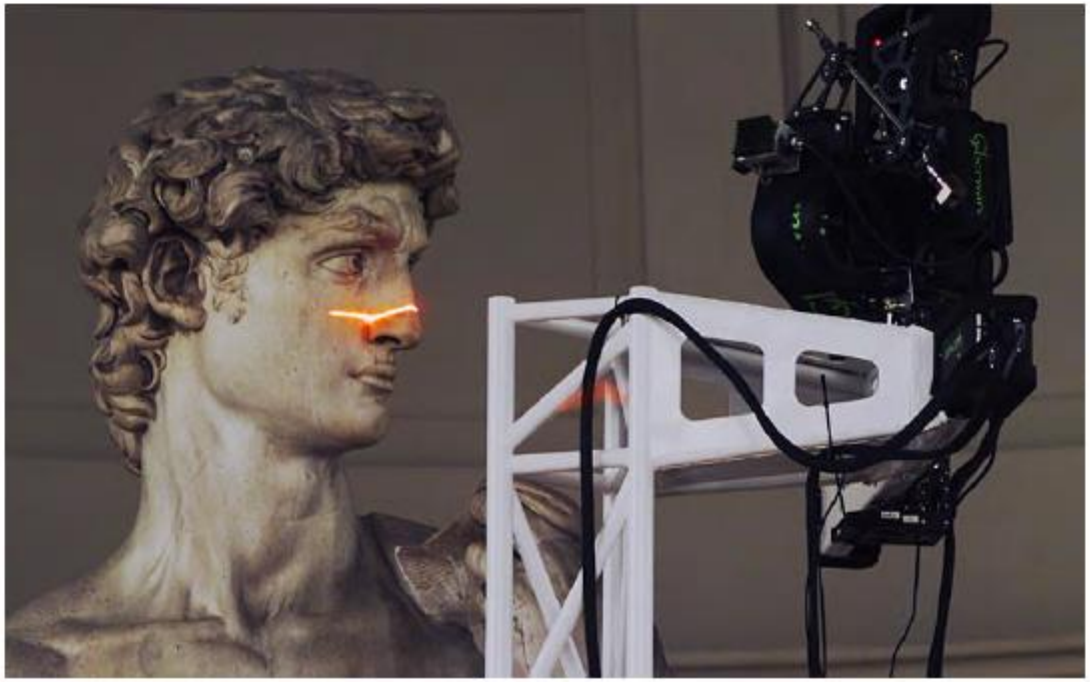
Other Stereo Pairs?

Mirrors



Taken from Mathieu and Devernay INRIA

Structure Light

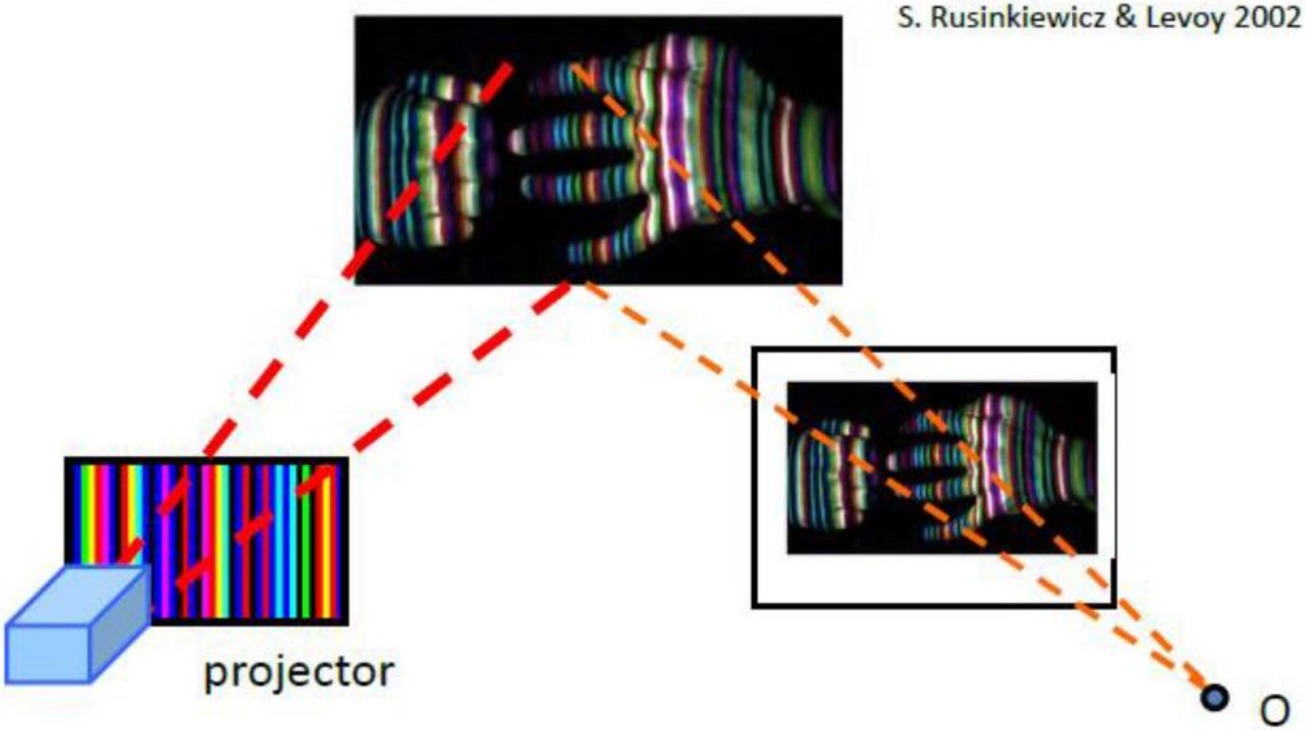


Digital Michelangelo Project
http://graphics.stanford.edu/projects/mich/

Structured Light

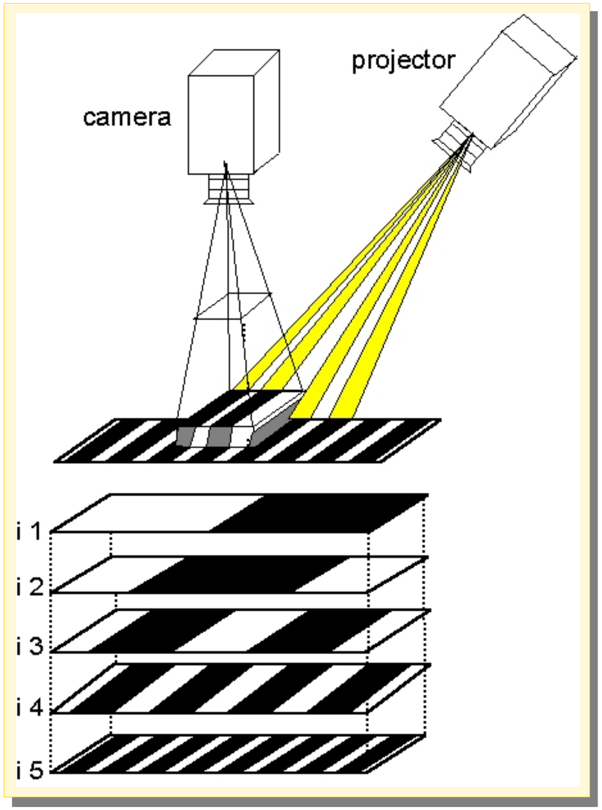
L. Zhang, B. Curless, and S. M. Seitz 2002

S. Rusinkiewicz & Levoy 2002

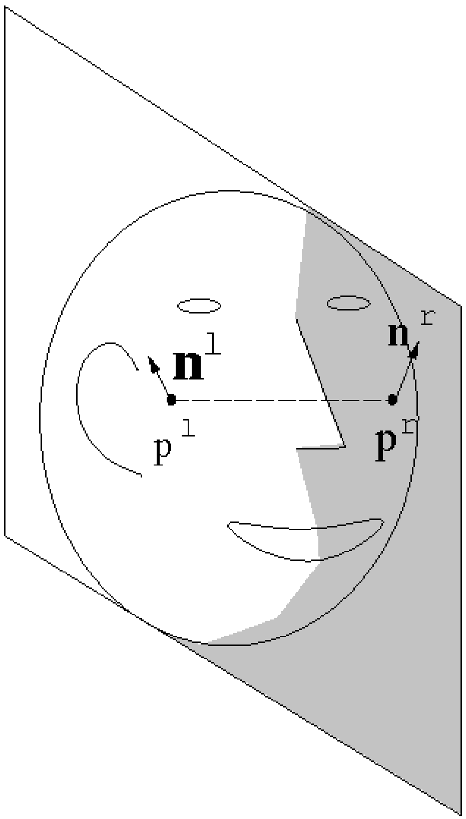


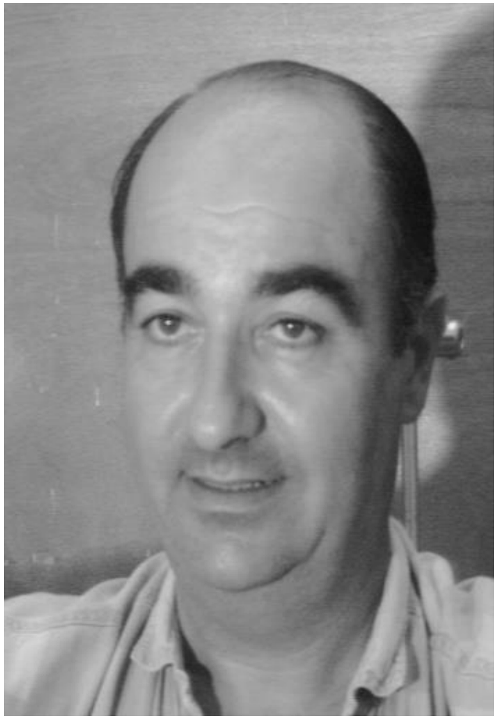
Structured Light

- known camera and projector geometry
- Depth can be recovered by triangulation



Bilateral Symmetry





So far

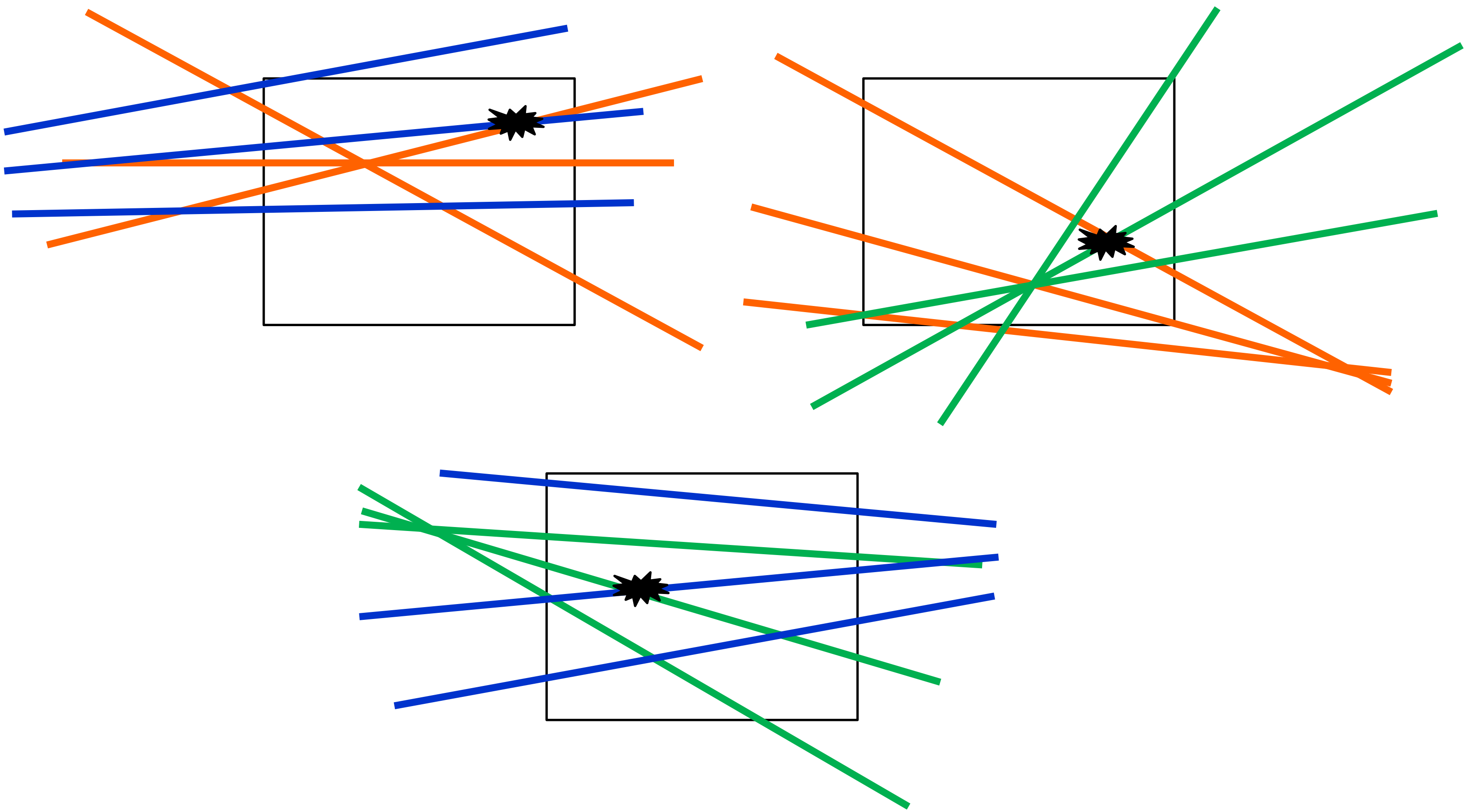
- Using algebra to compute:
  - Projection matrix
  - Triangulation
  - Calibration
  - Epipolar geometry
  - Rectification
- Correspondence: heuristics

הבינתחומי הרצליה
IDC HERZLIYA

Geometry

- Two images:
  - Stereo
  - Homography
- More images:
  - Multi-view stereo: improve correspondence
  - Structure from Motion: more views but uncalibrated

Improve Correspondence

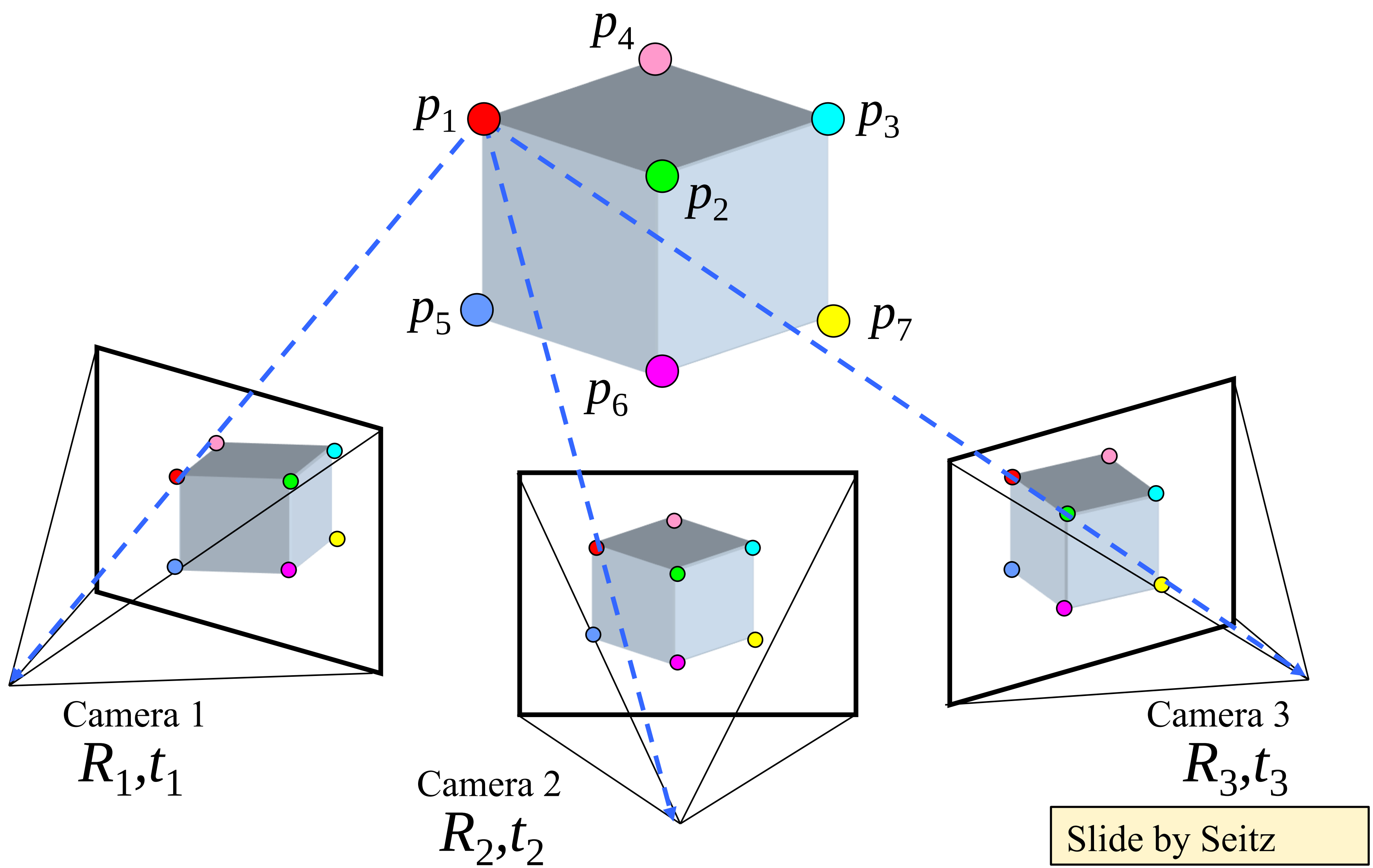


Structure from Motion (SfM)

- Given a sequence (or a set) of images compute:
  - The structure
  - The motion (camera location)




Structure from motion

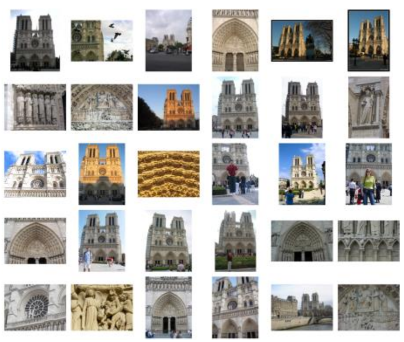


Slide by Seitz

Structure from Motion (SFM)

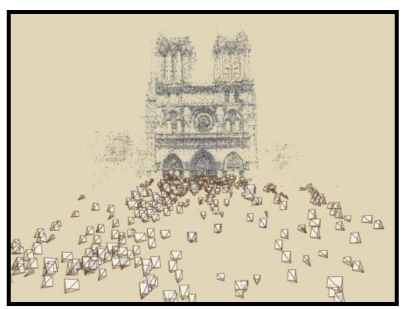
- Assumptions:
  - A set of images
  - Uncalibrated
- Applications:
  - Recover the camera locations
  - Photo Tourism
  - Improve robustness

Application: Photo Tourism Overview



Input photographs

Scene reconstruction



Relative camera positions and orientations

Point cloud

Sparse correspondence

Photo Explorer

Photo Tourism: Exploring Photo Collections in 3D, Snavely, Seitz, Szeliski

Slide by Seitz

SFM

- Given a set of m images with n corresponding points $\{\tilde{p}_{ij}\}$:
  - $\{M_i\}$ the set of unknown camera projections
  - $\{P_i\}$ the set of unknown 3D points
  - $\tilde{p}_{ij} = M_i \tilde{P}_j$
- Goal: find $\{M_i\}$ and $\{P_j\}$

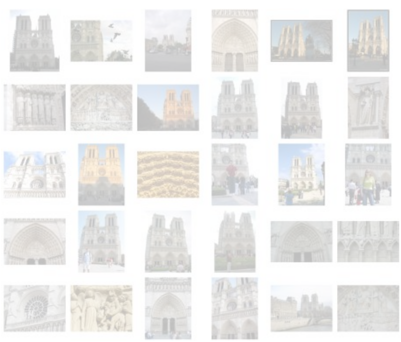
IDC HERZLIYA הבינתחומי הרצליה

SFM - Observations

- $\tilde{p}_{ij} = M_i \tilde{P}_j$
- Given $\{P_j\}$ we can compute $\{M_i\}$
- Given $\{M_i\}$ we can compute $\{P_j\}$
- Ambiguity: if $\{P_j\}$ and $\{M_i\}$ is a solution, then for any full rank 4×4 array A, $\{A^{-1} P_j\}$ and $\{M_i A\}$ are also a solution
  - $\tilde{p}_{ij} = M_i \tilde{P}_j = M_i A A^{-1} \tilde{P}_j$

IDC HERZLIYA
הבינתחומי הרצליה

Photo Tourism Overview



Input photographs

Scene reconstruction

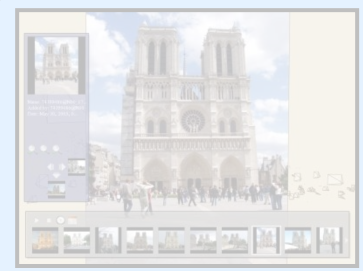
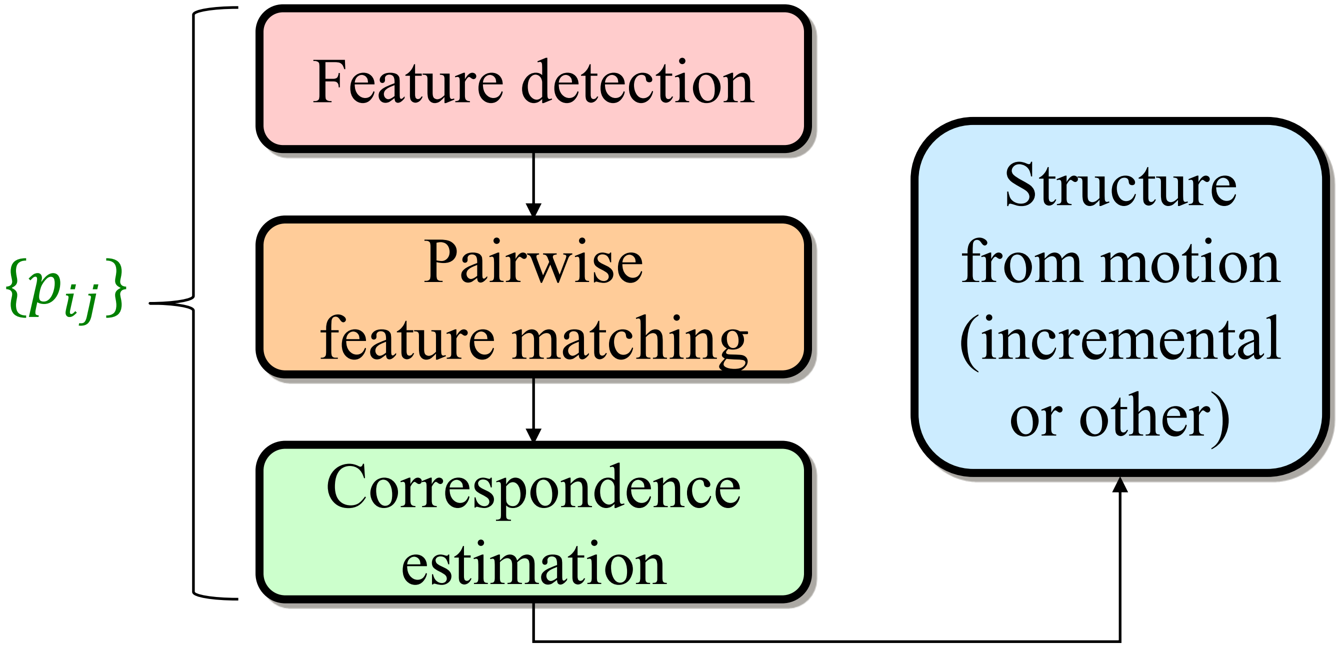


Photo Explorer

Slide by Seitz

Scene Reconstruction

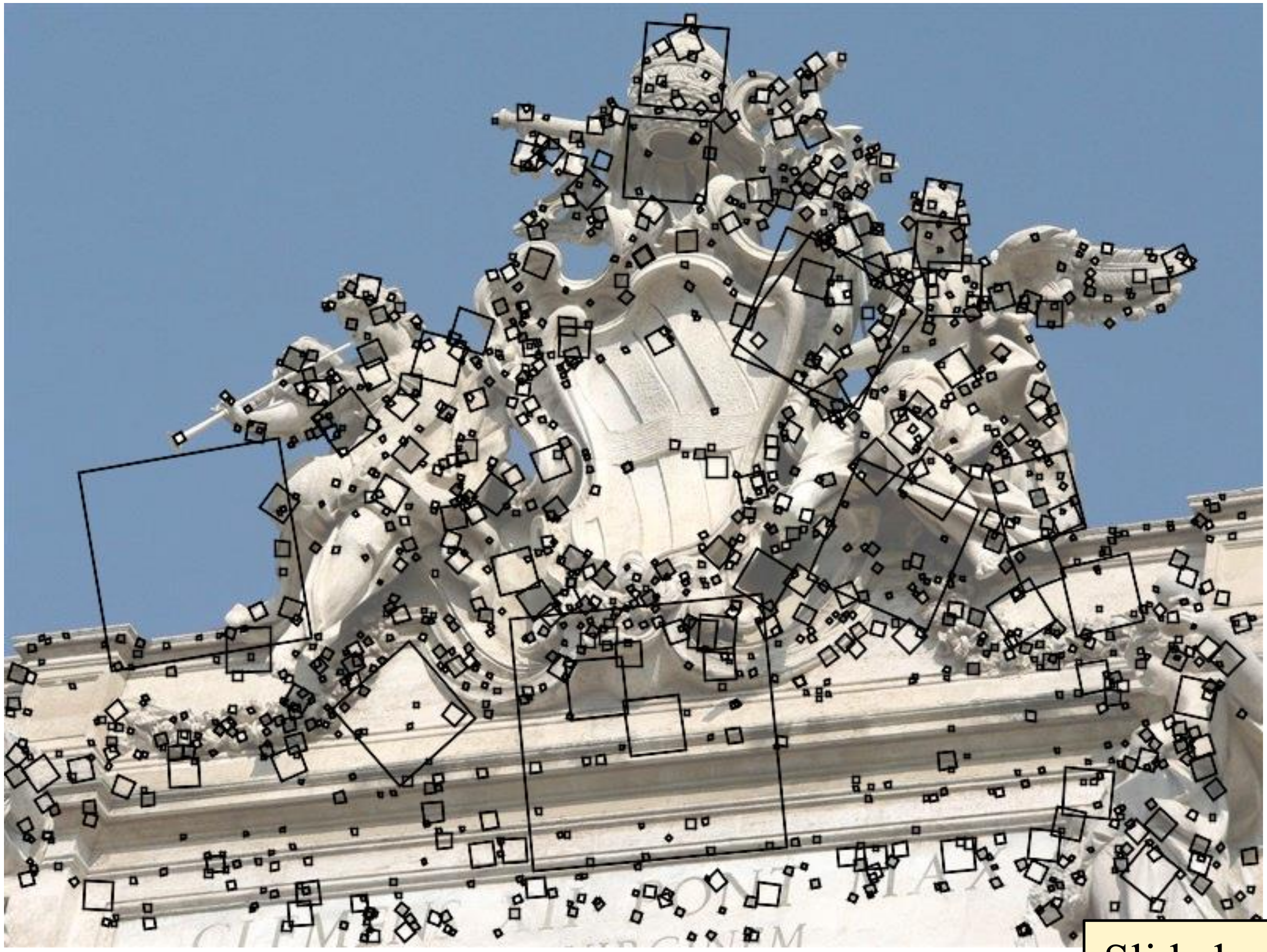
- Estimate $\{M_i\}$ and $\{P_i\}$



Slide by Seitz

Feature Detection

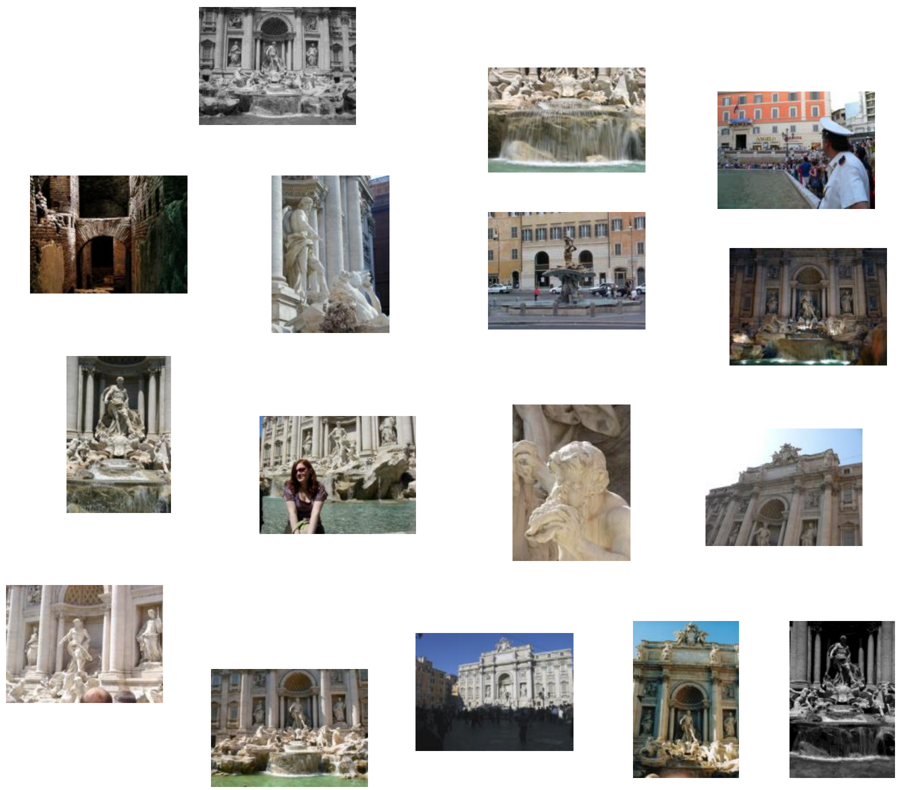
Detect features using SIFT [Lowe, IJCV 2004]



Slide by Seitz

Feature Detection

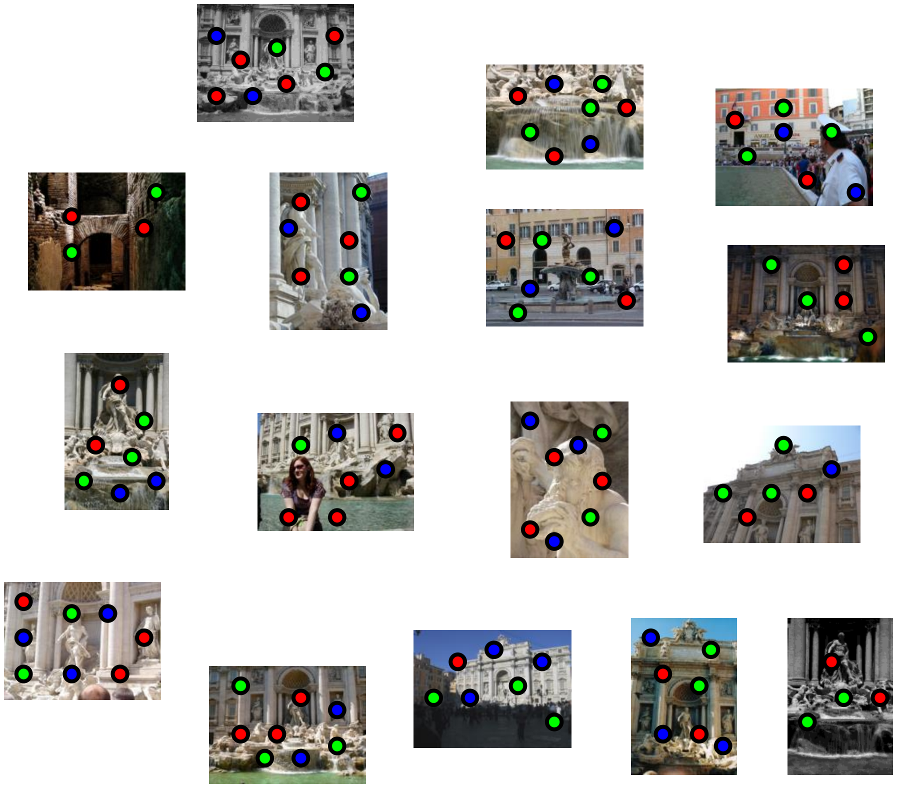
Detect features using SIFT [Lowe, IJCV 2004]



Slide by Seitz

Feature Detection

Detect features using SIFT [Lowe, IJCV 2004]

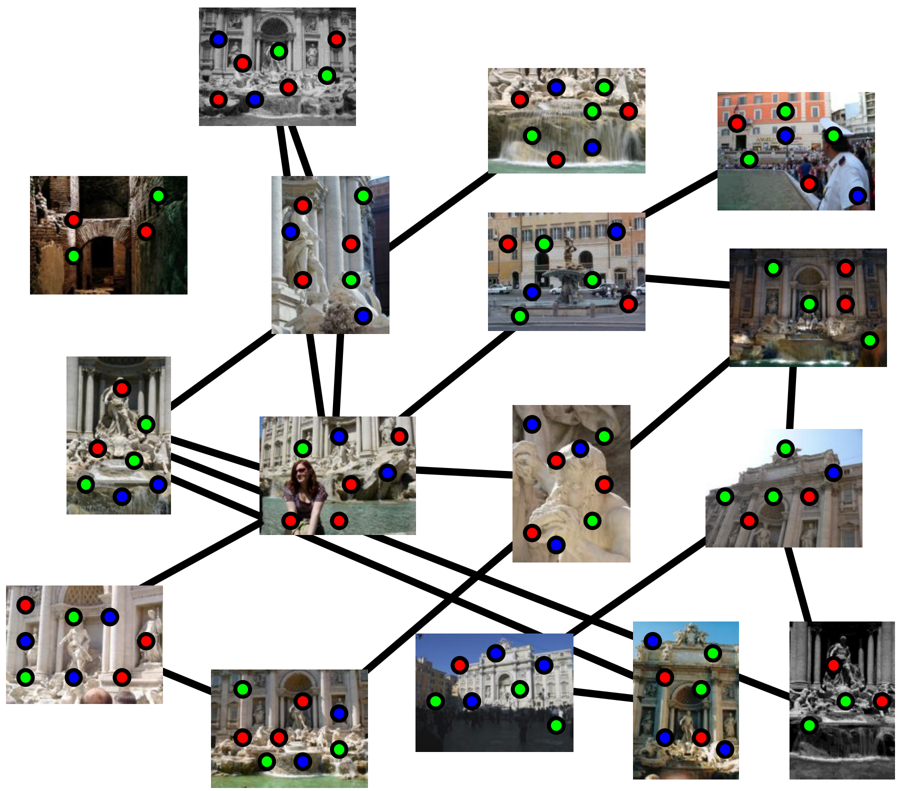


Slide by Seitz

IDC HERZLIYA הבינתחומי הרצליה

Feature Matching

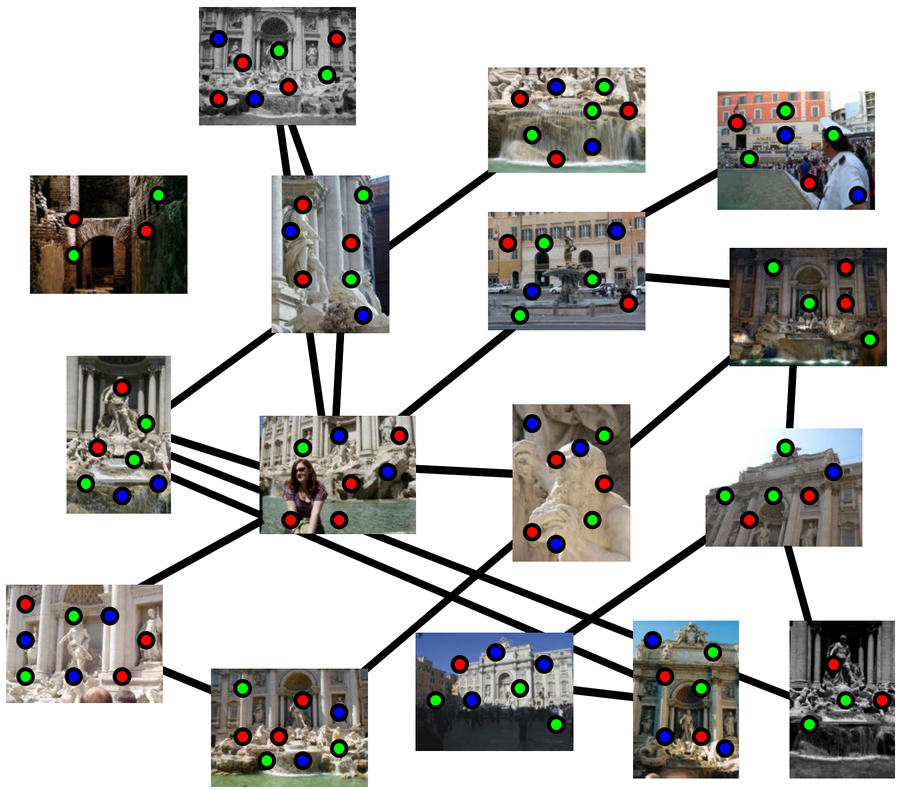
Match features between each pair of images



Slide by Seitz

Feature matching

Refine matching using RANSAC [Fischler & Bolles 1987] to estimate fundamental matrices between pairs



Slide by Seitz

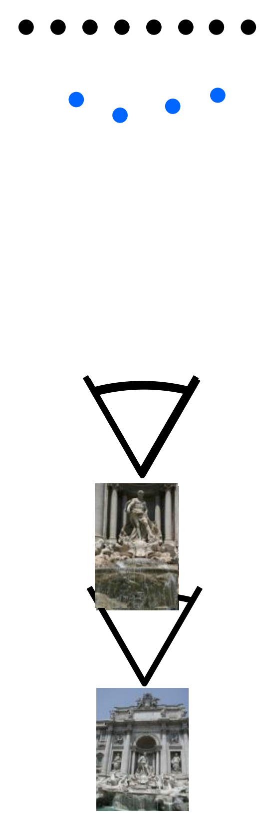
Incremental SFM

Trevi Fountain,
Rome

Given $\{P_i\}$ compute $\{M_i\}$
Given $\{M_i\}$ compute $\{P_i\}$

הבינתחומי הרצליה
IDC HERZLIYA

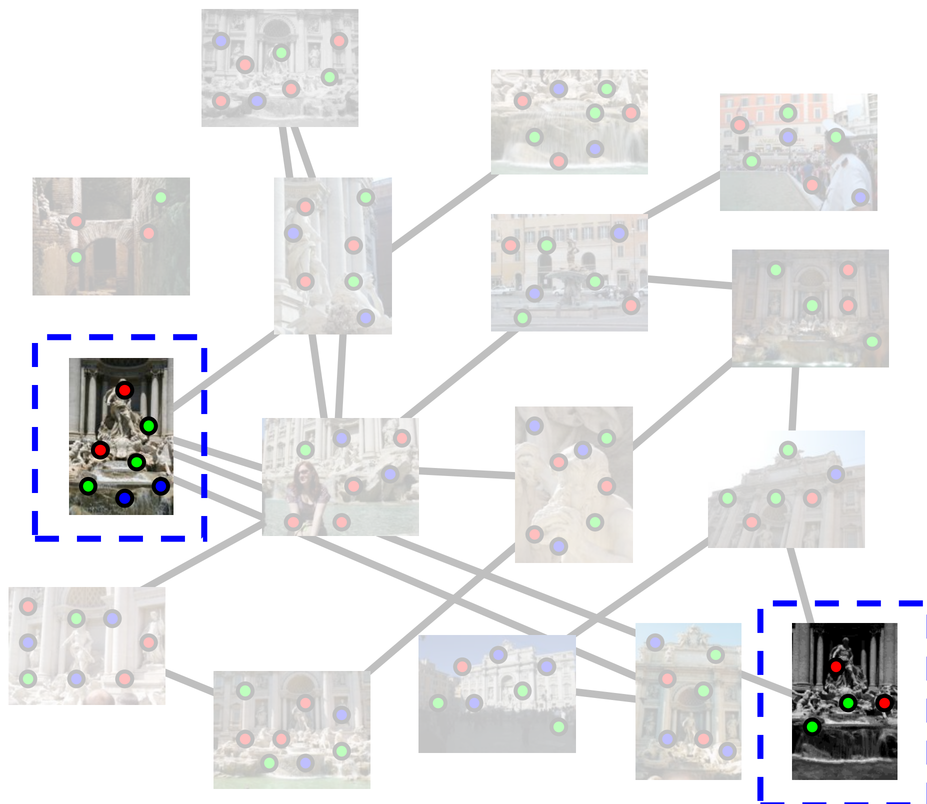
Incremental SFM



Given $\{P_i\}$ compute $\{M_i\}$

Given $\{M_i\}$ compute $\{P_i\}$

Perspective Projection: Incremental SFM



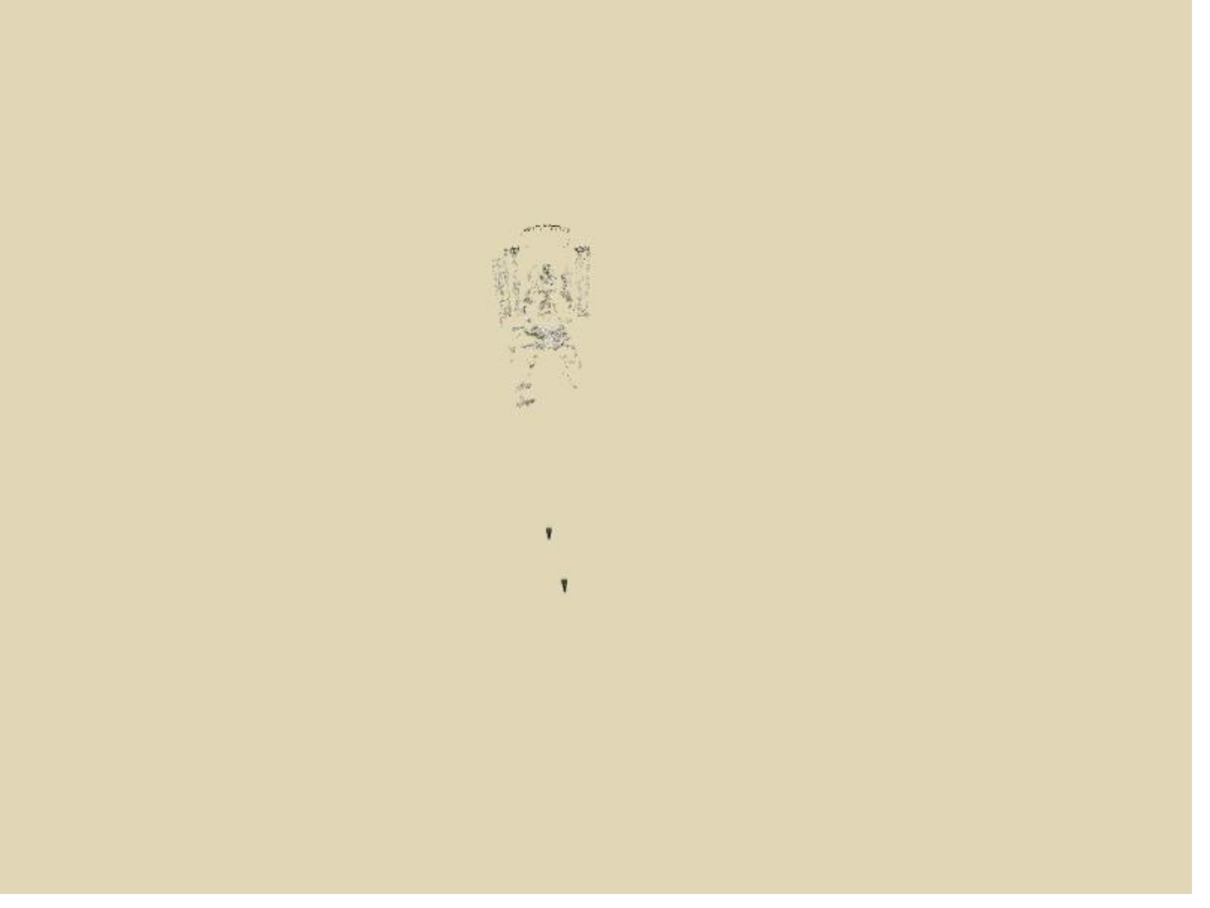
Given $\{P_i\}$ compute $\{M_i\}$

Given $\{M_i\}$ compute $\{P_i\}$

Incremental SFM

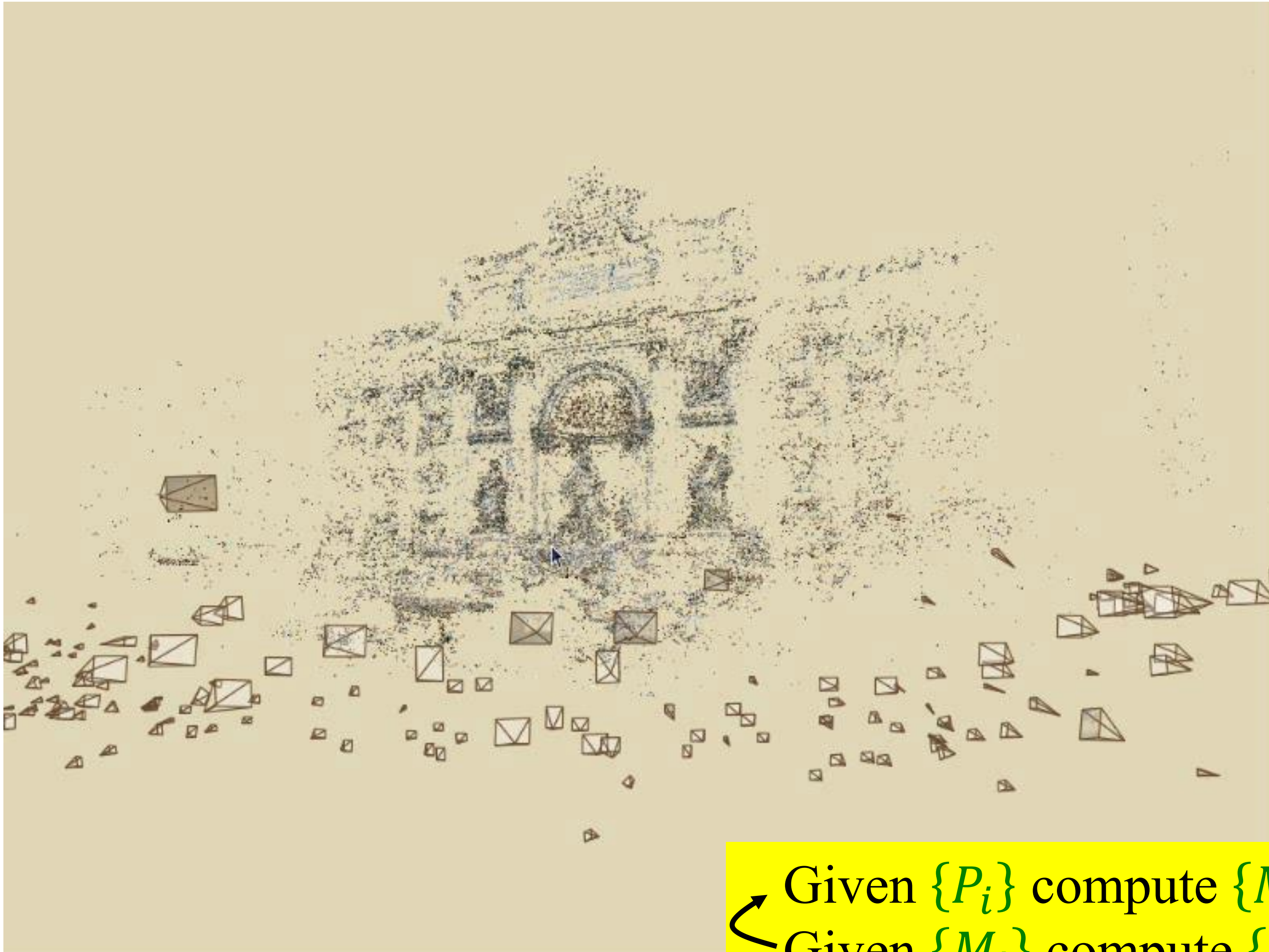
Given $\{P_i\}$ compute $\{M_i\}$
Given $\{M_i\}$ compute $\{P_i\}$

Incremental SFM



Given $\{P_i\}$ compute $\{M_i\}$
Given $\{M_i\}$ compute $\{P_i\}$

Photo Explorer



Given $\{P_i\}$ compute $\{M_i\}$

Given $\{M_i\}$ compute $\{P_i\}$

Challenges & Limitations

- A heuristic algorithm- not necessarily an optimal solution
- Which order to use the images?
- How it affects the results?
- Efficiency

Affine Structure from Motion

The Problem: Reconstruct scene geometry and camera parameters from two or more images

Assumptions:

- Known correspondence
- Orthographic projection

Only approximation of perspective

Advantage: Closed form solution

Next Class

- SfM for orthographic projection
- Motion analysis
  - Change detection
  - Optical flow
  - Tracking