Multiview Alignment and Sparse SFM

CSE P576

Dr. Matthew Brown

Multiview + Sparse SFM

- Multiview Image Alignment, Residuals, Error Function
- Structure from Motion (SFM)
- Bundle Adjustment, Pose Estimation, Triangulation

[Szeliski 7, 9]

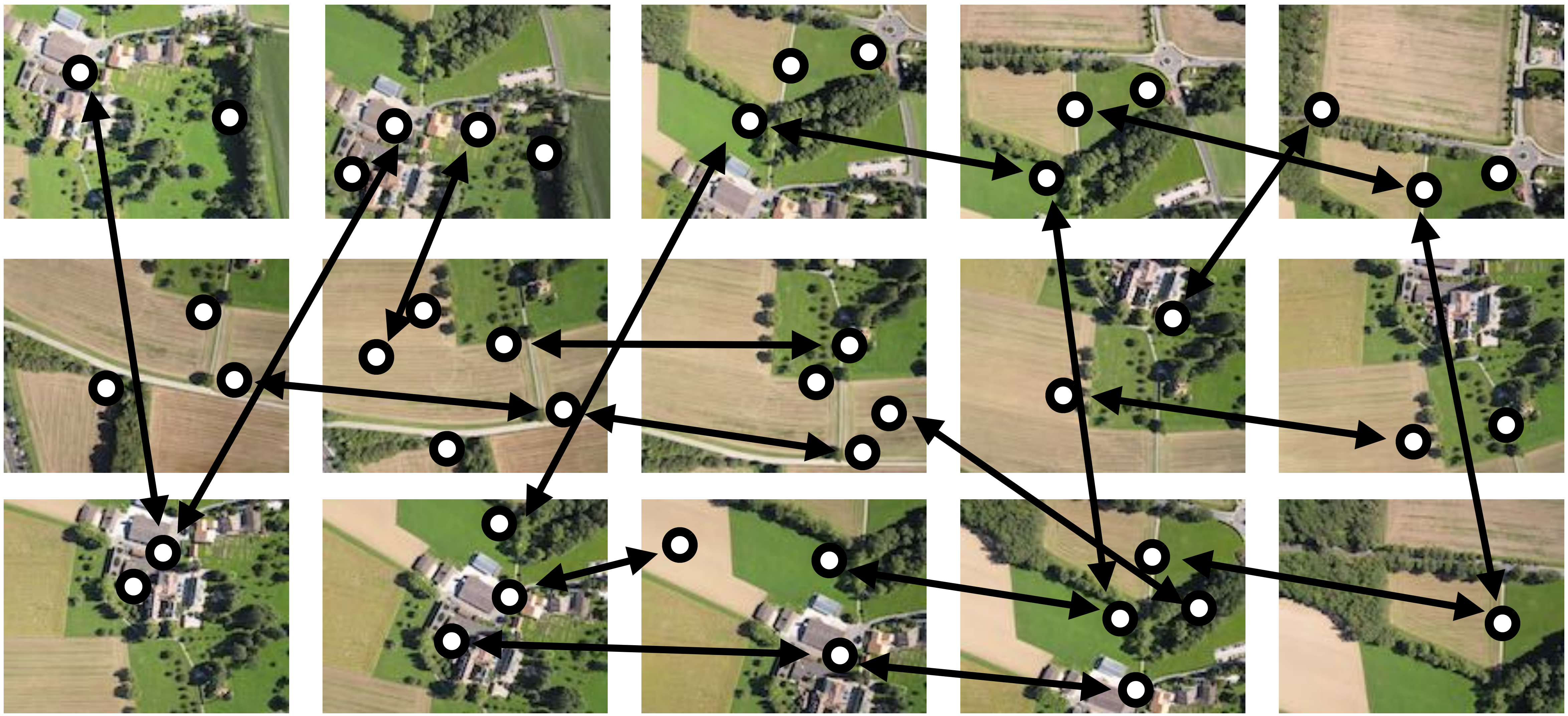
Multiview Image Alignment

- Align a set of images given a motion model (e.g., planar affine)

[Szeliski 9.2]

Multiview Image Alignment

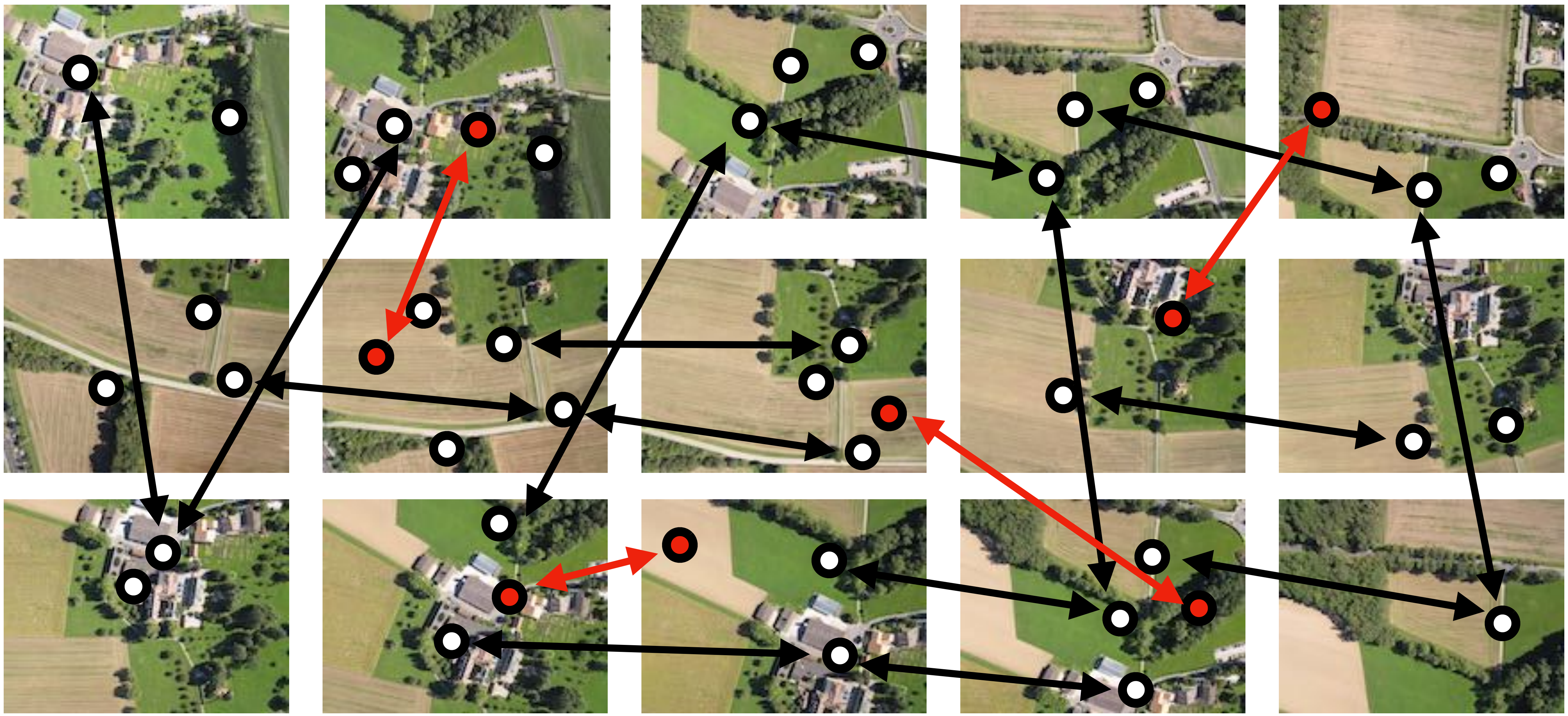
- Align a set of images given a motion model (e.g., planar affine)



Step 1: Find all matches between images using SIFT

Multiview Image Alignment

- Align a set of images given a motion model (e.g., planar affine)



Step 1: Find all matches between images using SIFT

Step 2: Remove incorrect matches using RANSAC

Multiview Image Alignment

- Align a set of images given a motion model (e.g., planar affine)

Step 1: Find all matches between images using SIFT

Step 2: Remove incorrect matches using RANSAC

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

4 inliers (red, yellow, orange, brown),
4 outliers (blue, light blue, purple, pink)

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

4 inliers (red, yellow, orange, brown),

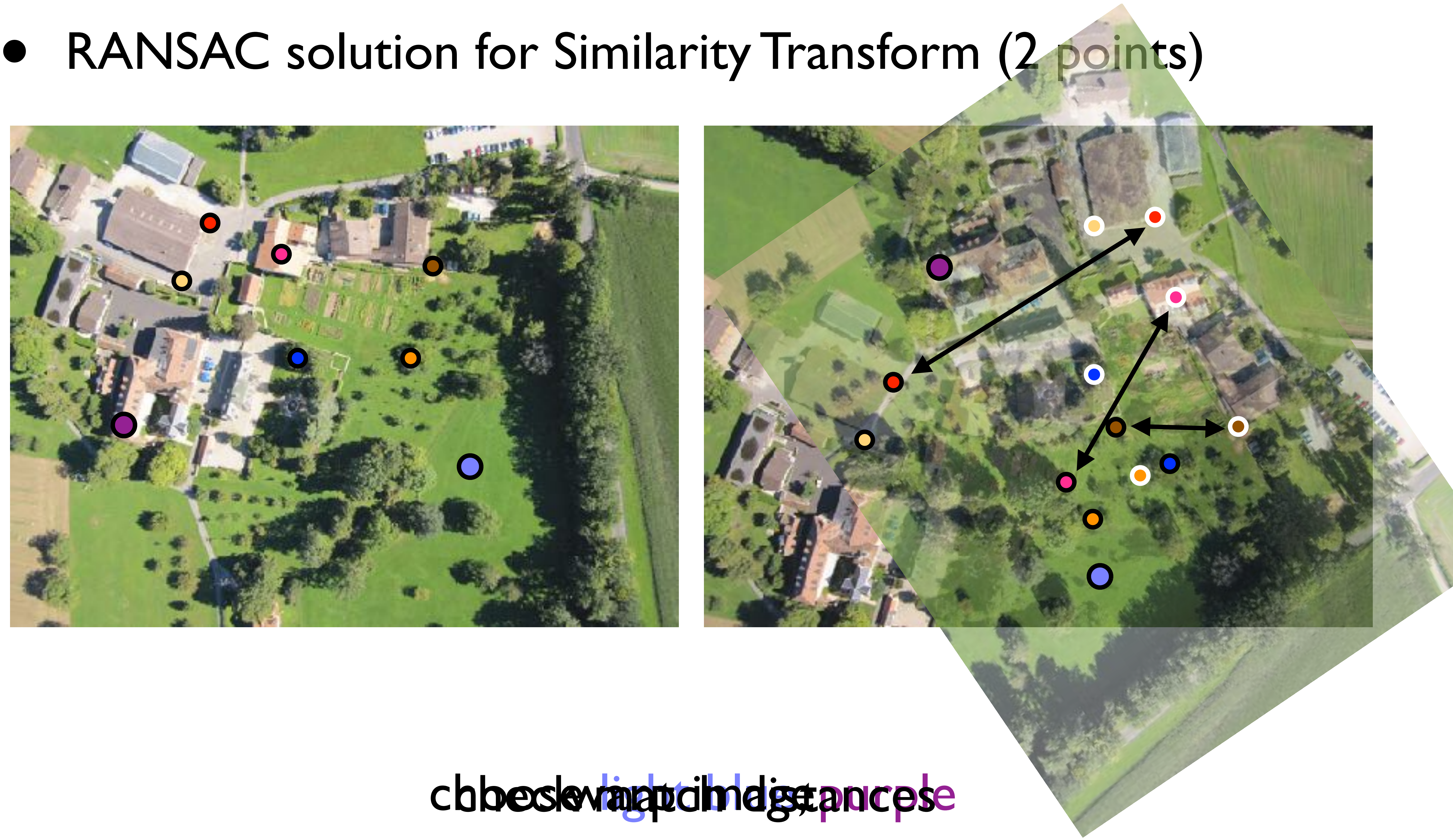
RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

4 outliers (blue, light blue, purple, pink)

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)



check match distances

light blue, purple

#inliers = 2

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

choose pink blue

warp image

check distances

#inliers = 2

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

choose red, orange

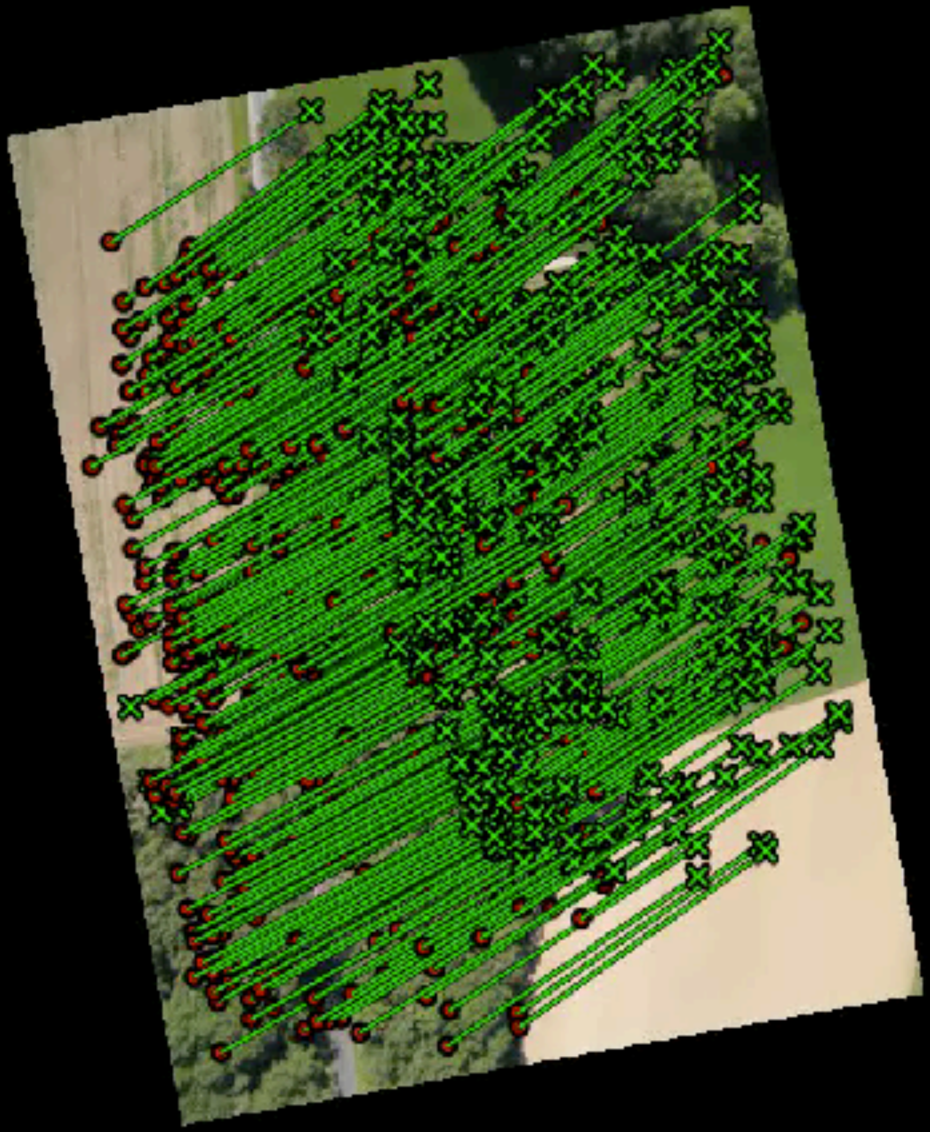
#inliers = 4

RANSAC recap

- RANSAC solution for Similarity Transform (2 points)

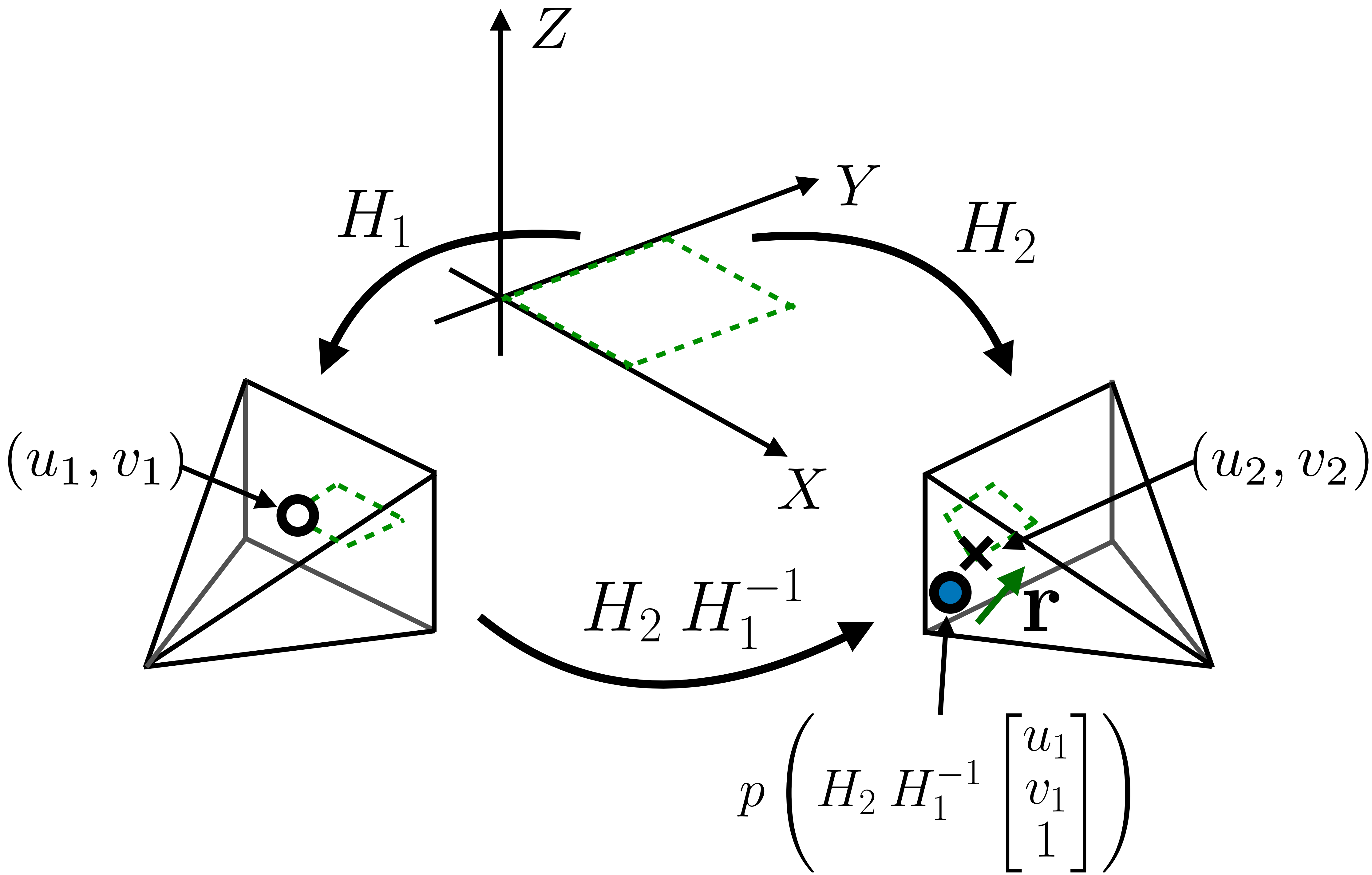
Planar Image Alignment

- Given a clean set of correspondences, align all images



Planar Mapping Residuals

- Residual = vector between observed feature and projection

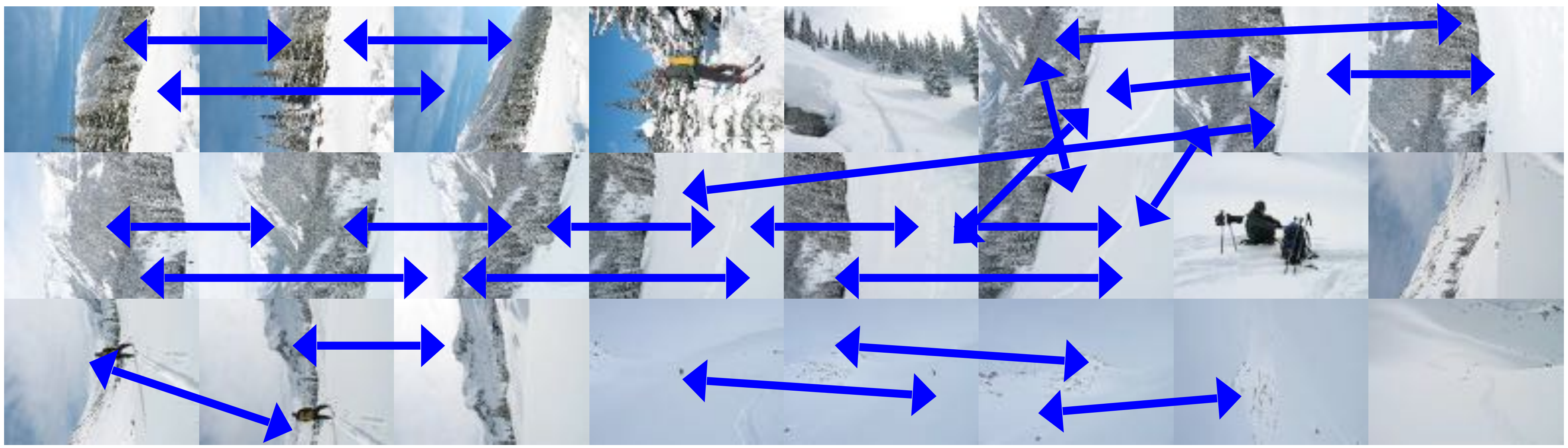


Multiview Image Alignment

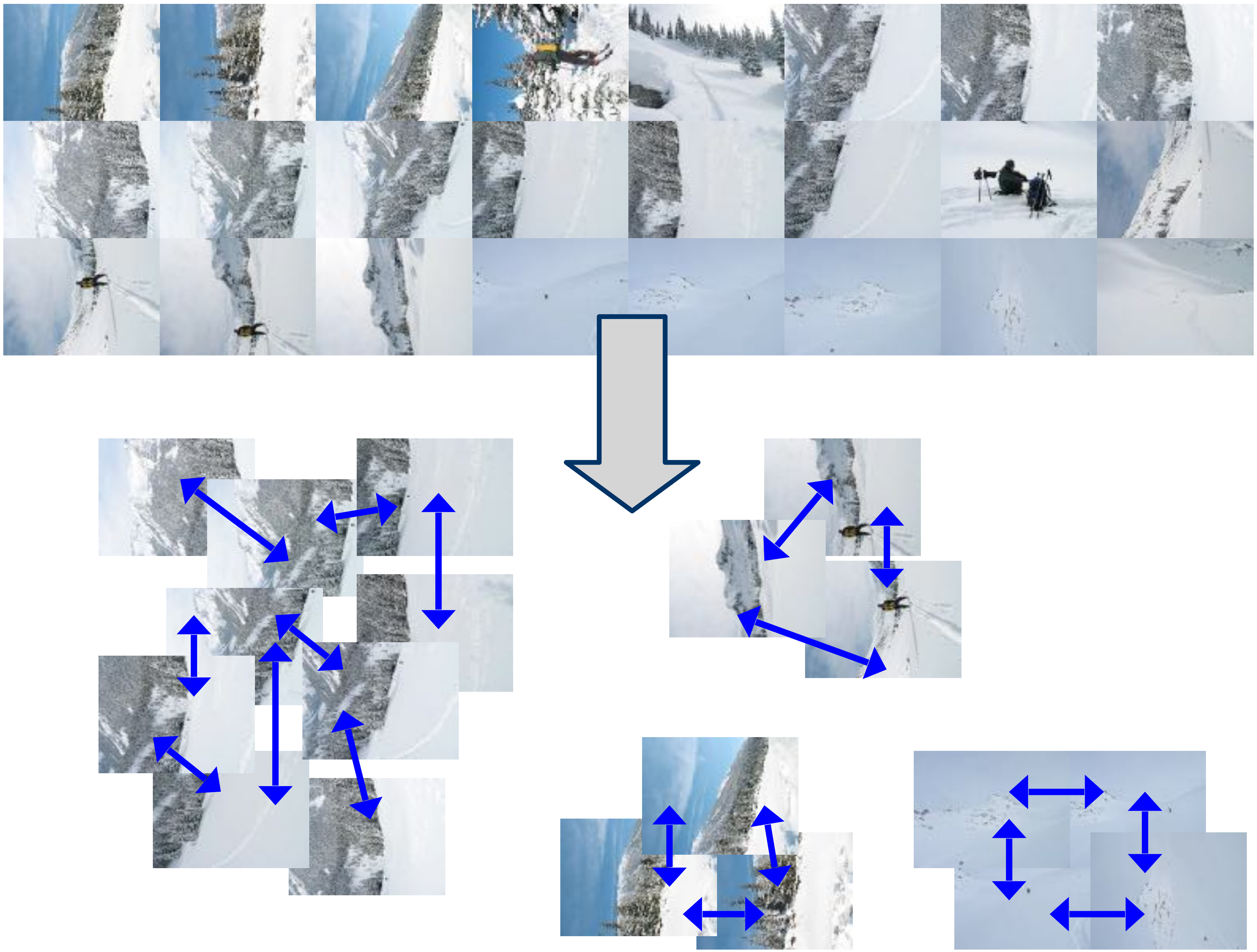
- Minimize squared projection errors between images with respect to planar transform parameters (H matrices)

 4.5

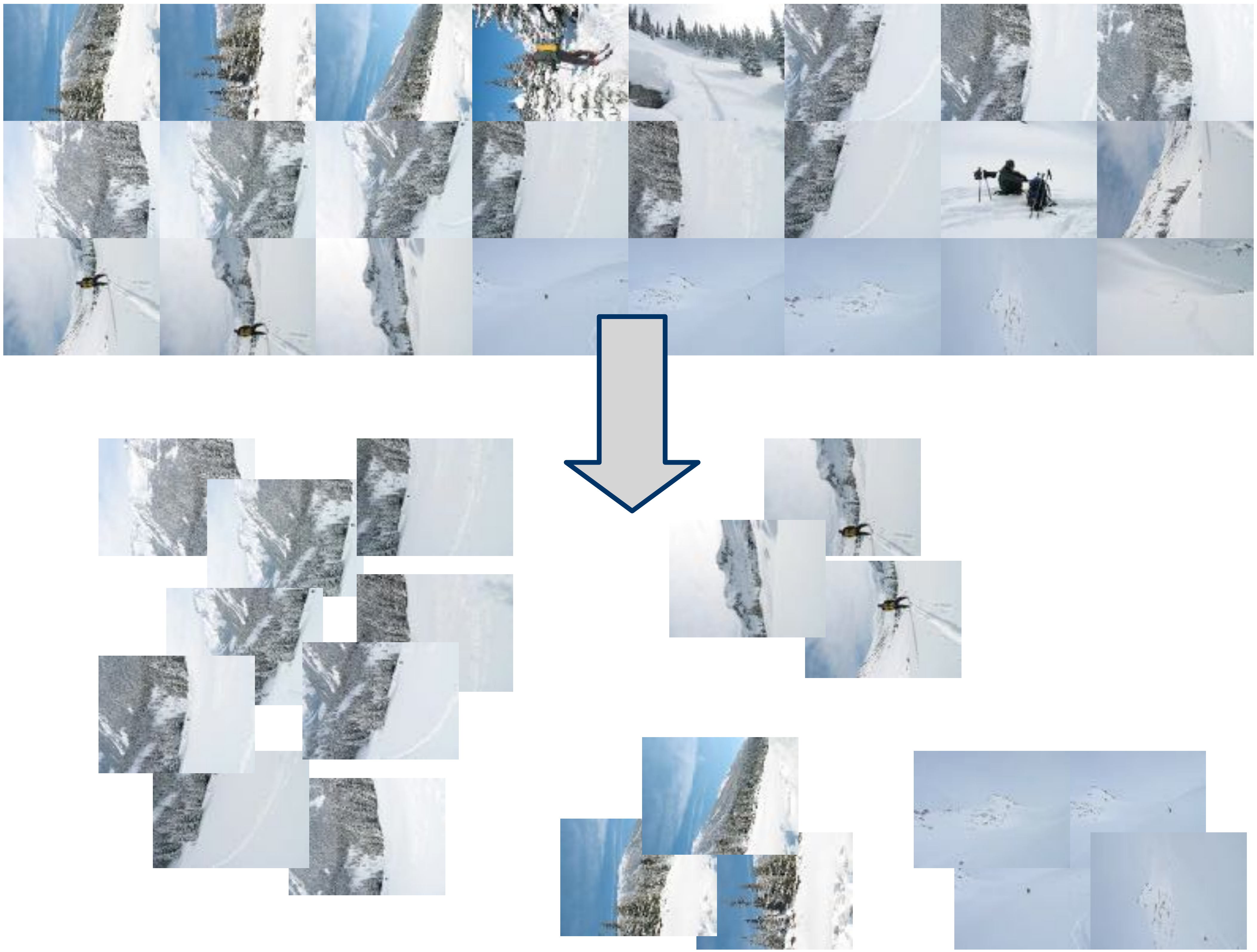
Panorama Recognition



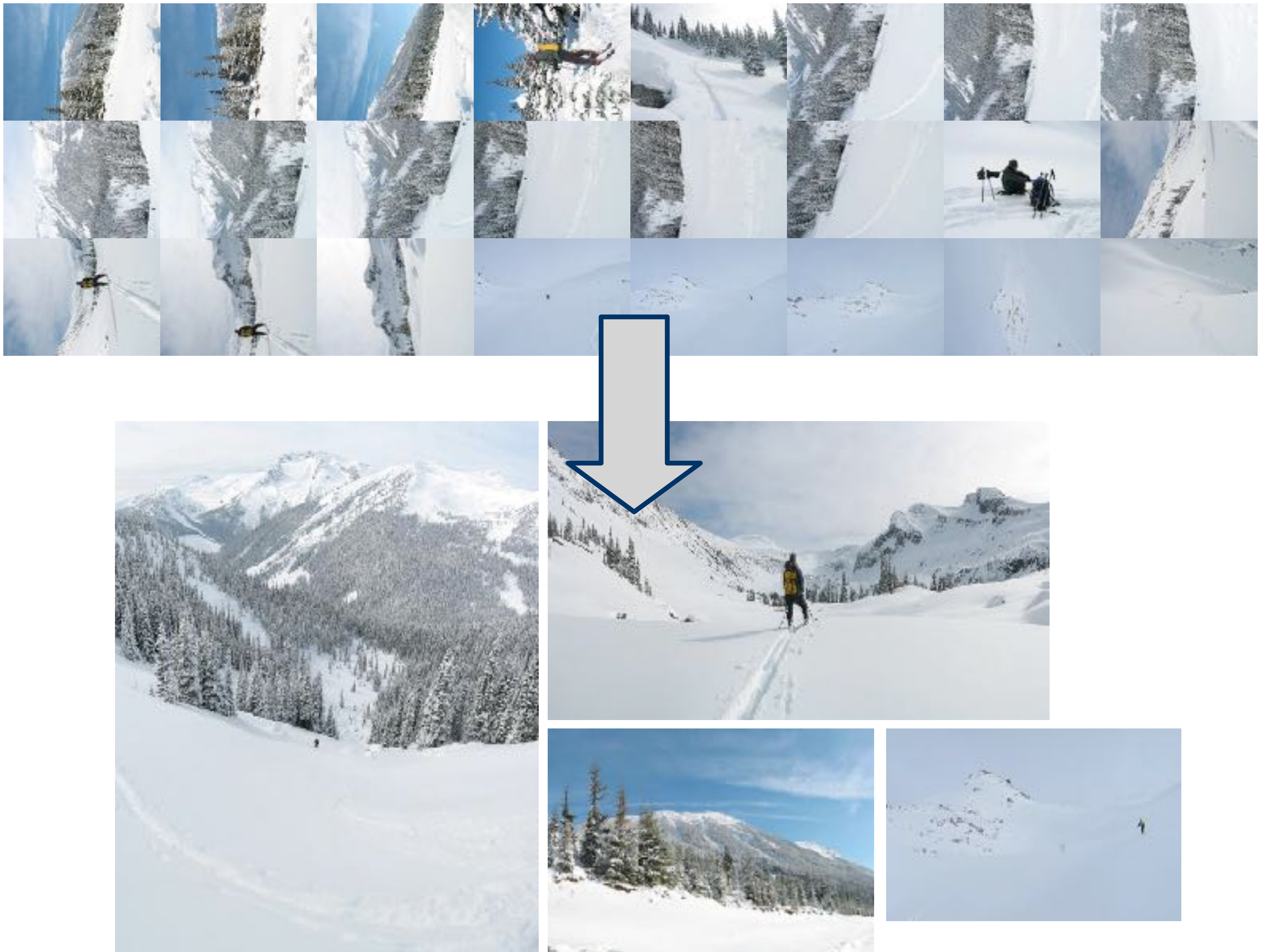
Panorama Recognition



Panorama Recognition



Panorama Recognition



Panorama Stitching

- We can concatenate pairwise homographies, but over time multiple pairwise mappings accumulate errors
- We use global alignment (bundle adjustment) to close the gap

Structure from Motion

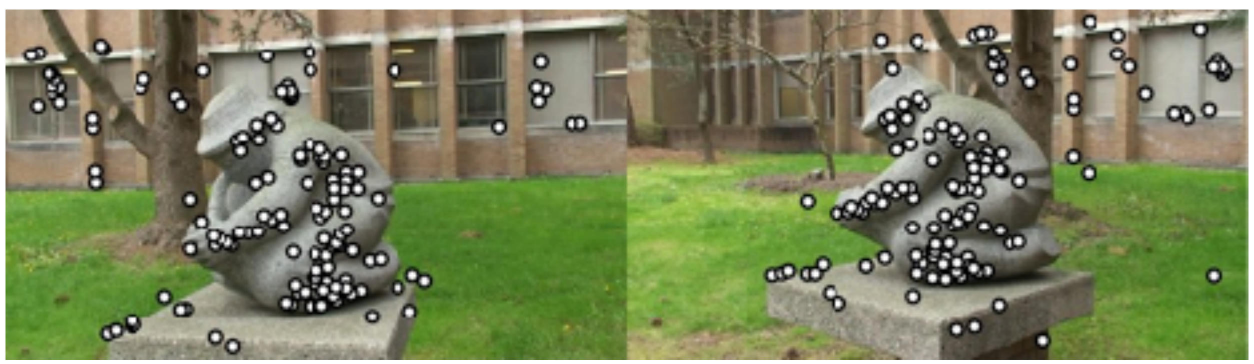
Given an (unordered) set of input images, compute cameras and 3D structure of the scene

[Szeliski 7]

Structure from Motion

2-view Structure from Motion

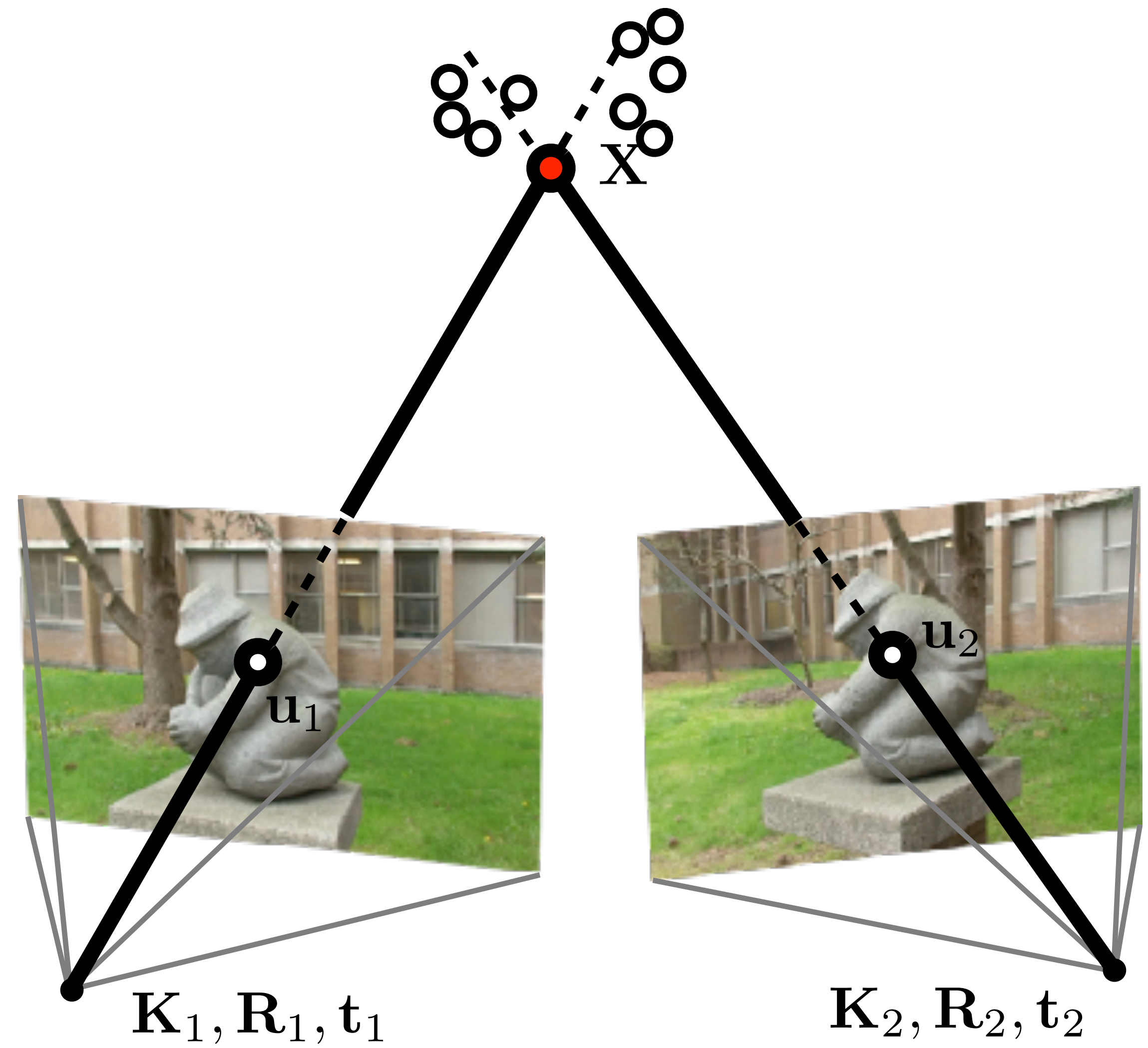
- We can use the combination of SIFT/RANSAC and triangulation to compute 3D structure from 2 views



Raw SIFT matches

RANSAC for F

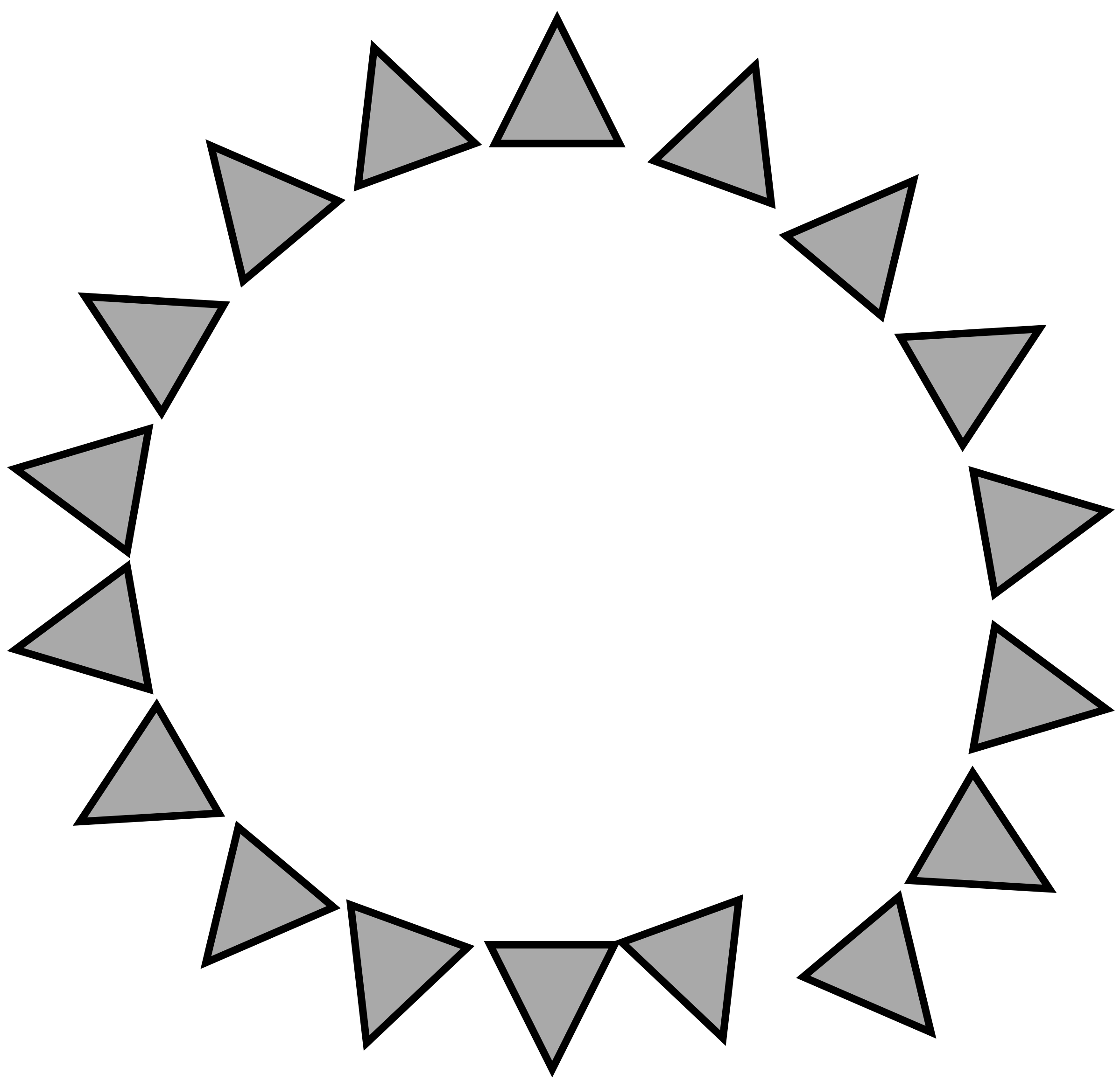
Extract R, t



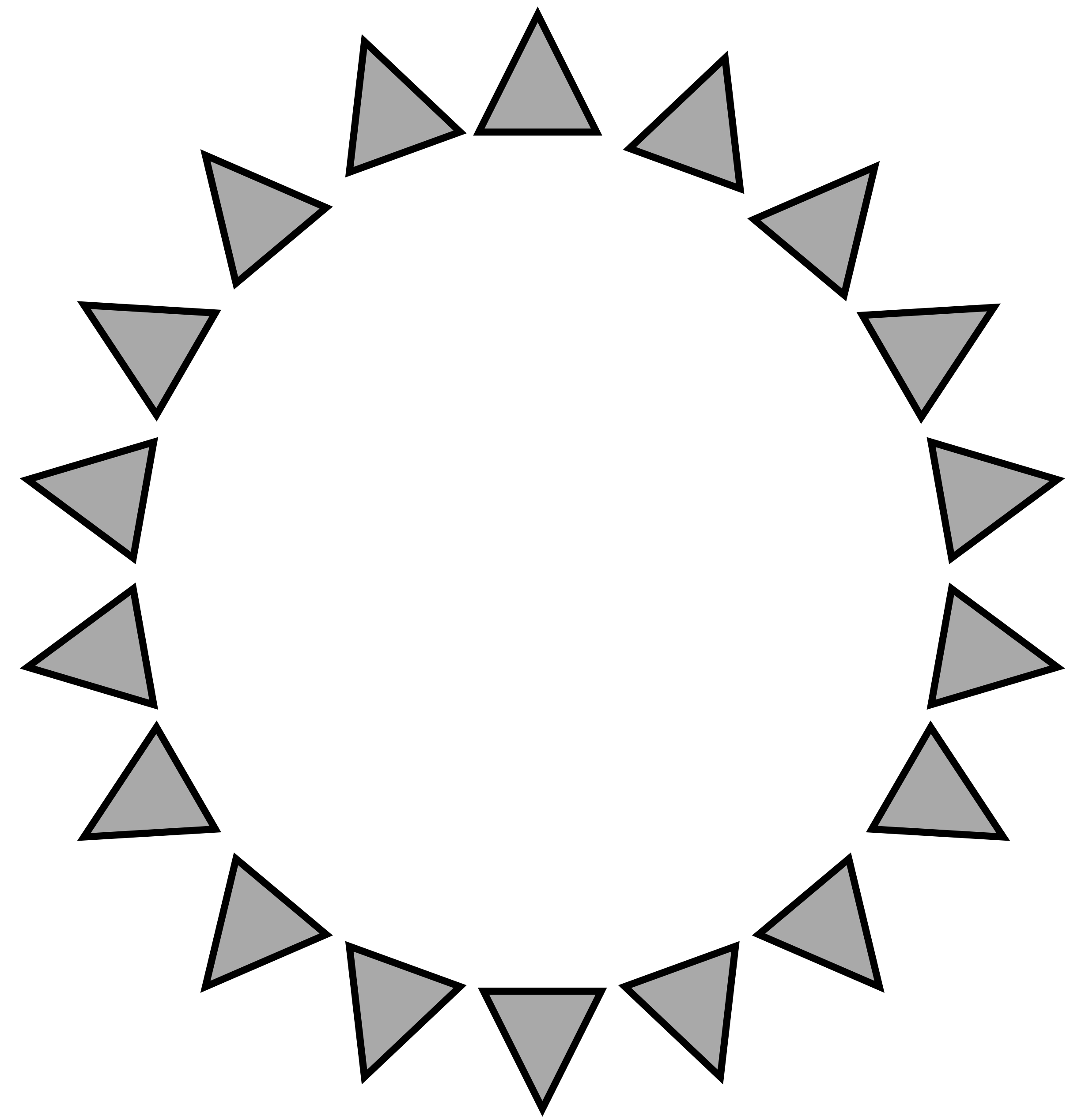
Triangulate to 3D Point Cloud

Global Alignment

- Concatenation of pairwise R, t estimates results in drift, e.g.,



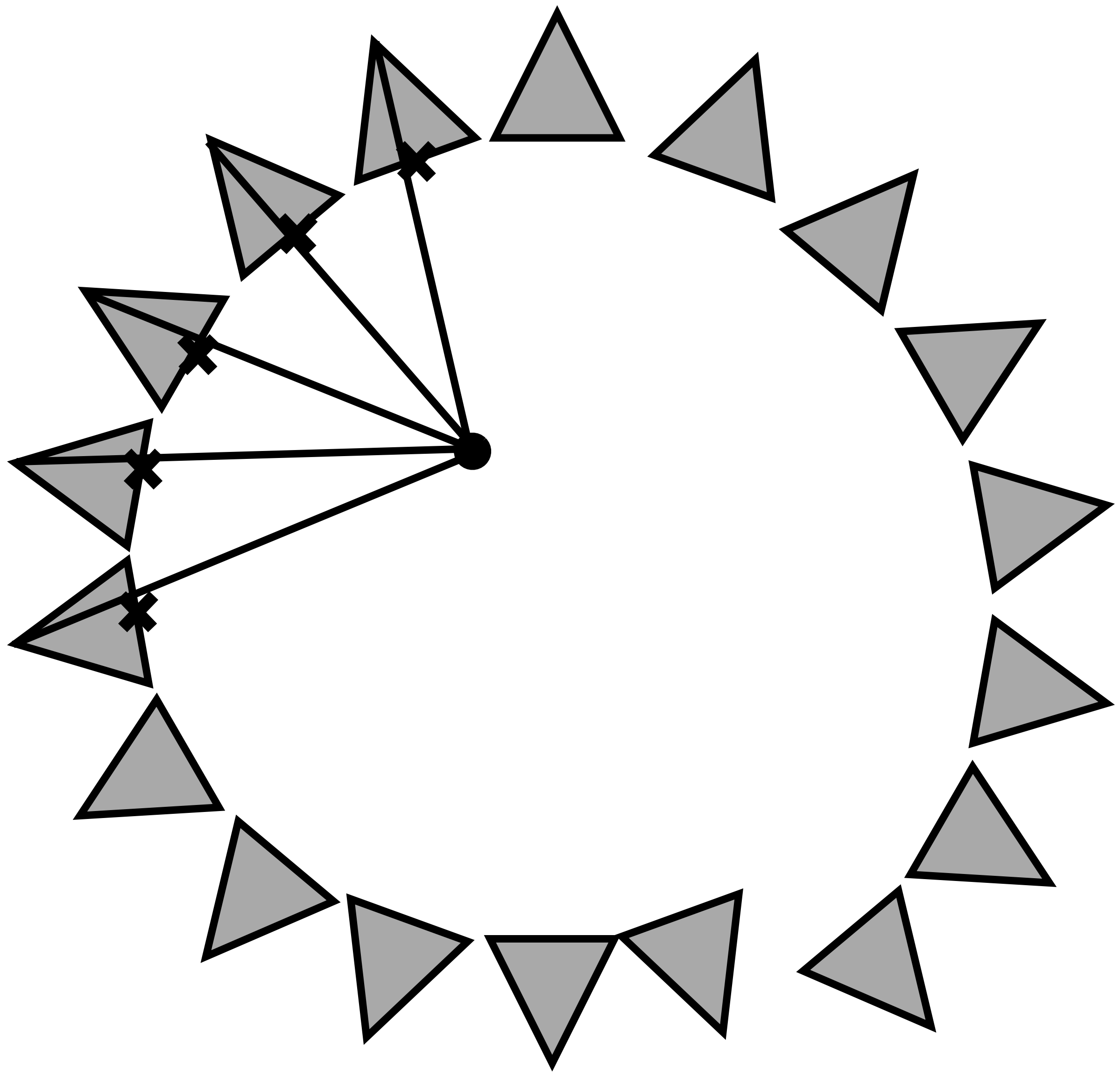
Pairwise alignment



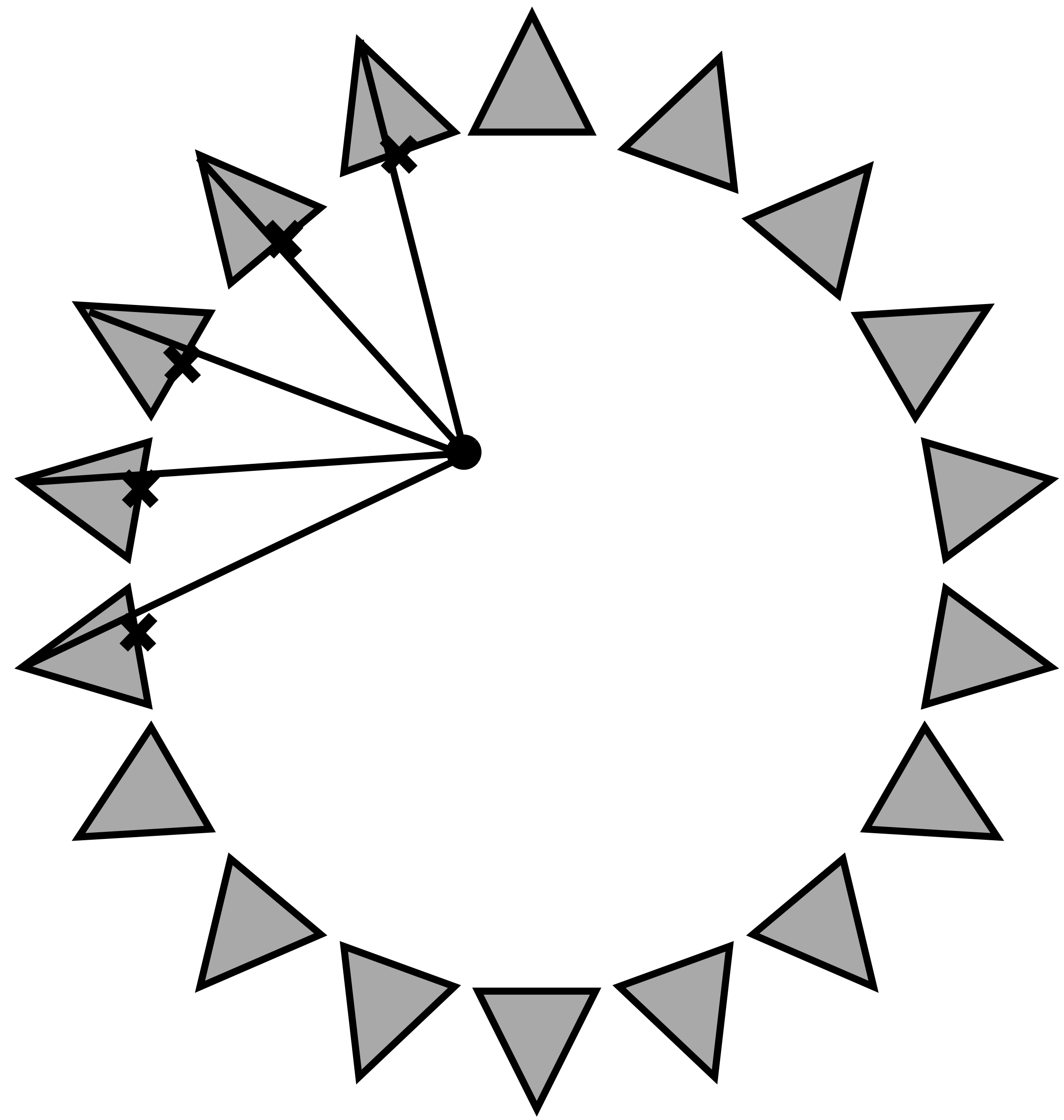
Global alignment

Global Alignment

- Concatenation of pairwise R, t estimates results in drift, e.g.,



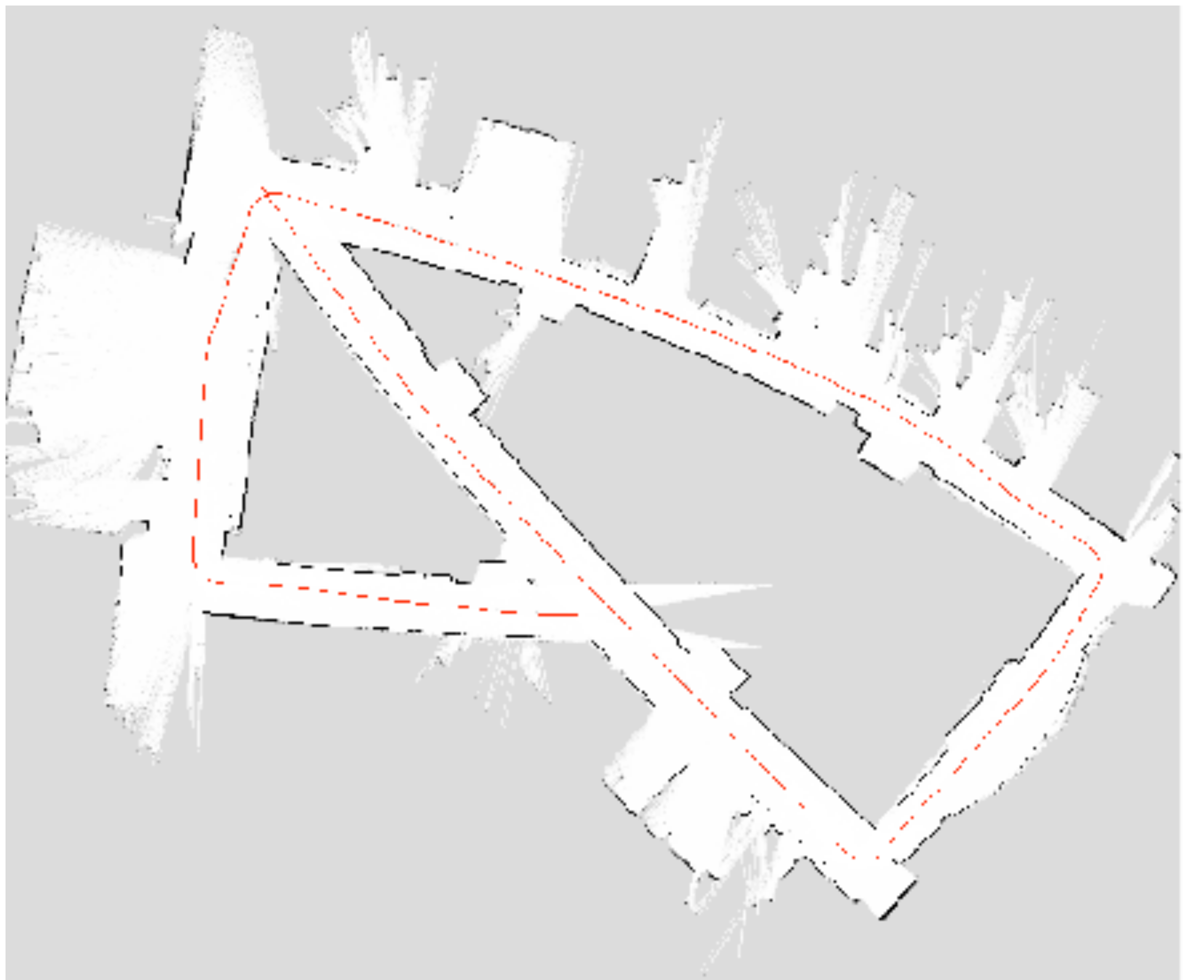
Pairwise alignment



Global alignment

Global Alignment

- In robotic navigation frame-frame alignment also causes drift



We can use **bundle adjustment** to close the gap

[Kaess Dellaert 2010]

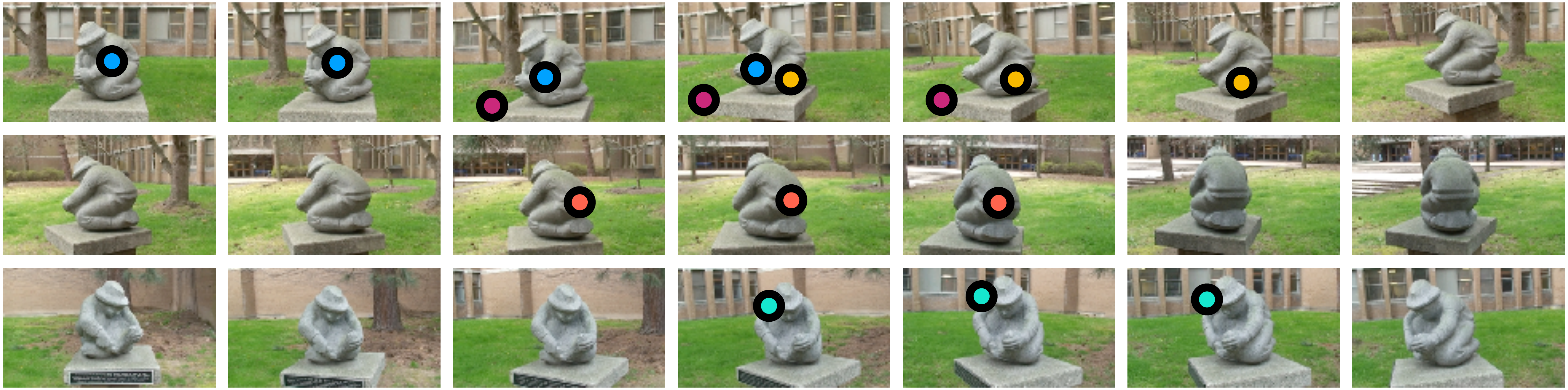
RANSAC for F

Raw feature matches (after ratio test filtering)

Solved for F and RANSAC inliers

Feature Tracking

- Form feature tracks by combining pairwise feature matches

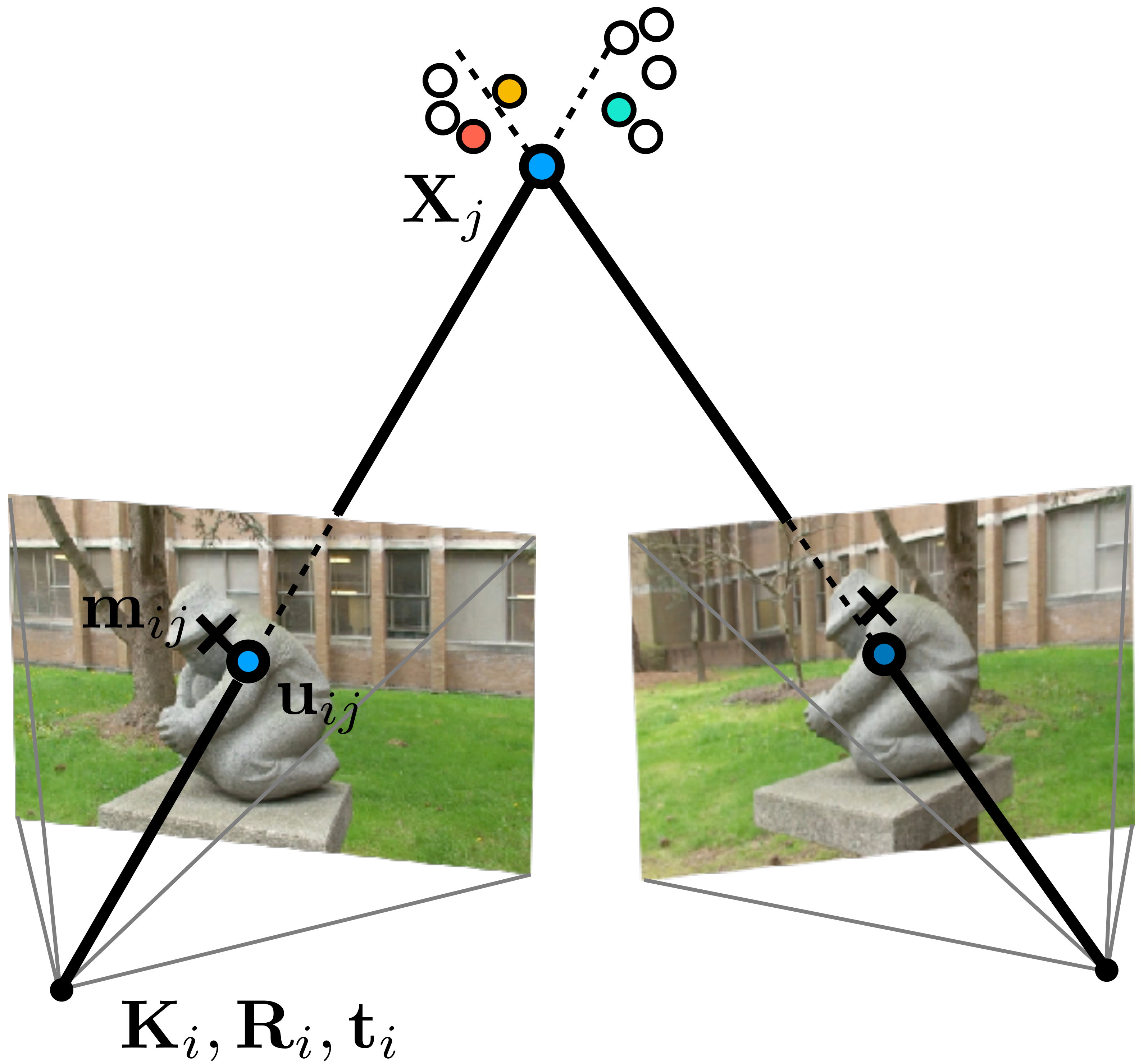


- Tracked features become individual 3D points in the reconstruction
- Features matched across 3 or more views provide strong constraints on the 3D reconstruction

Bundle Adjustment

- Minimise errors projecting 3D points into all images

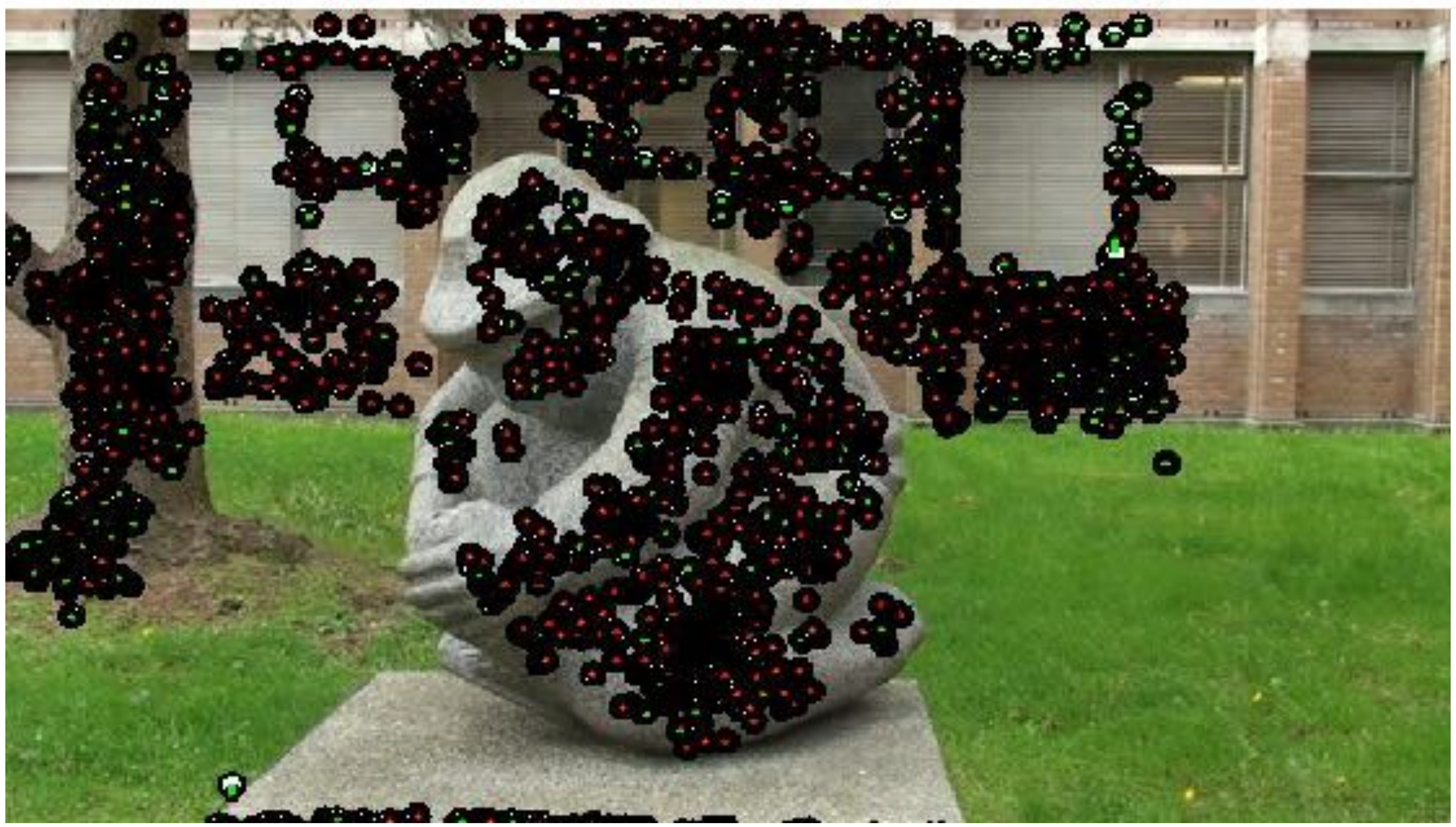
4.6



[Szeliski 7.4]

Bundle Adjustment

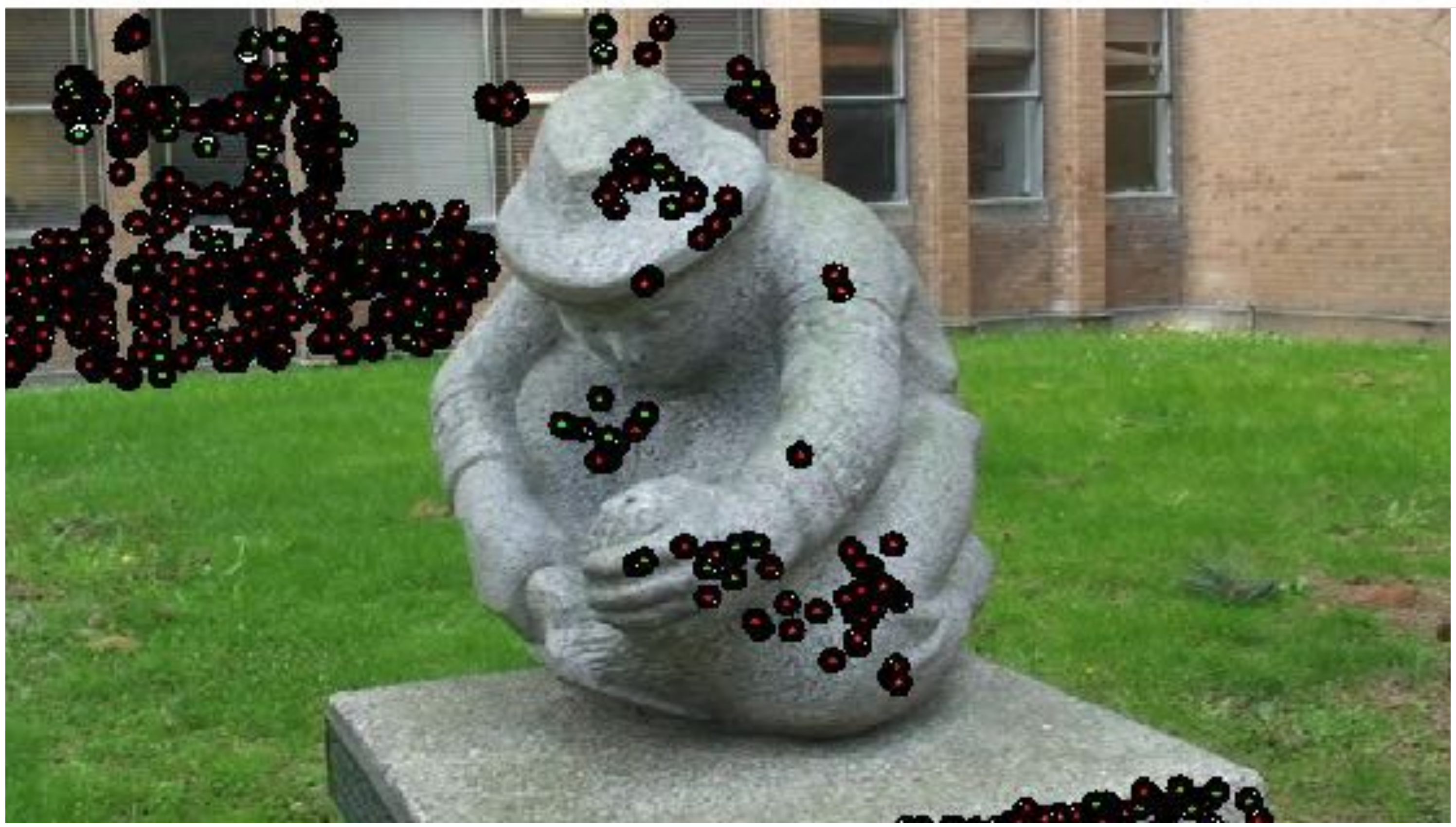
- Initialization with 3 views



Joint optimization of cameras and structure

Bundle Adjustment

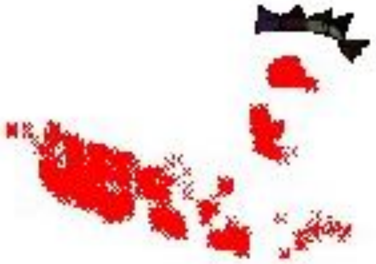
- Add camera 4



Estimate camera pose, add new 3D points, jointly optimize

Bundle Adjustment

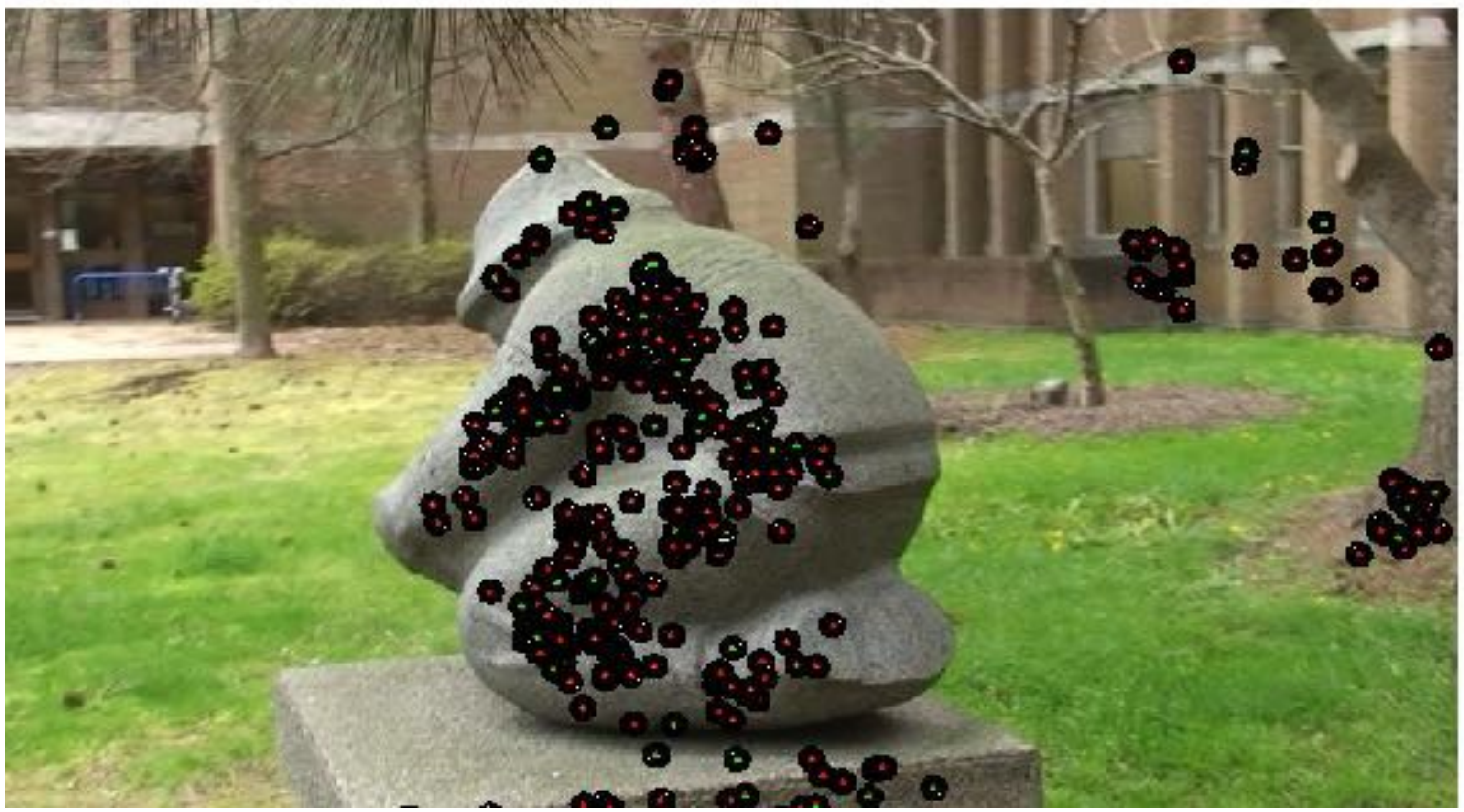
- Add camera 5

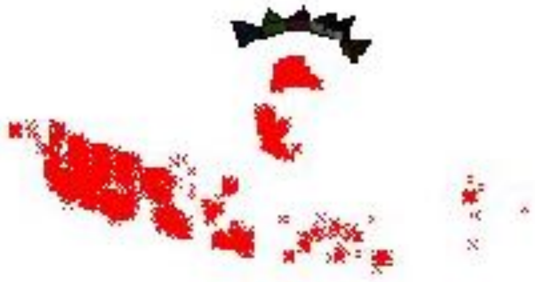


Estimate camera pose, add new 3D points, jointly optimize

Bundle Adjustment

- Add camera 6

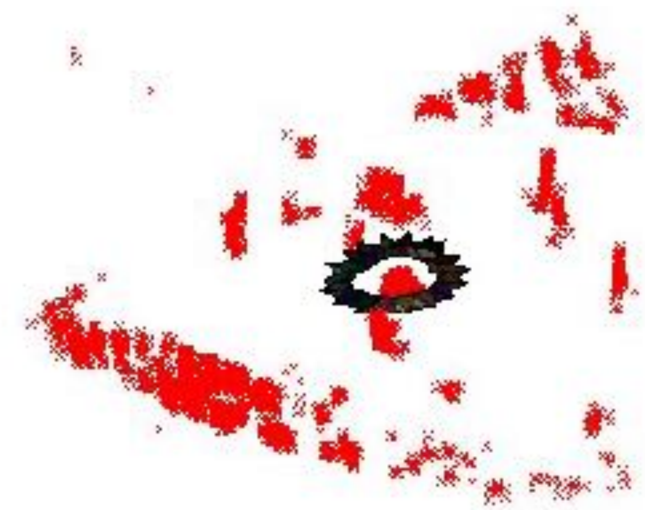




Estimate camera pose, add new 3D points, jointly optimize

Bundle Adjustment

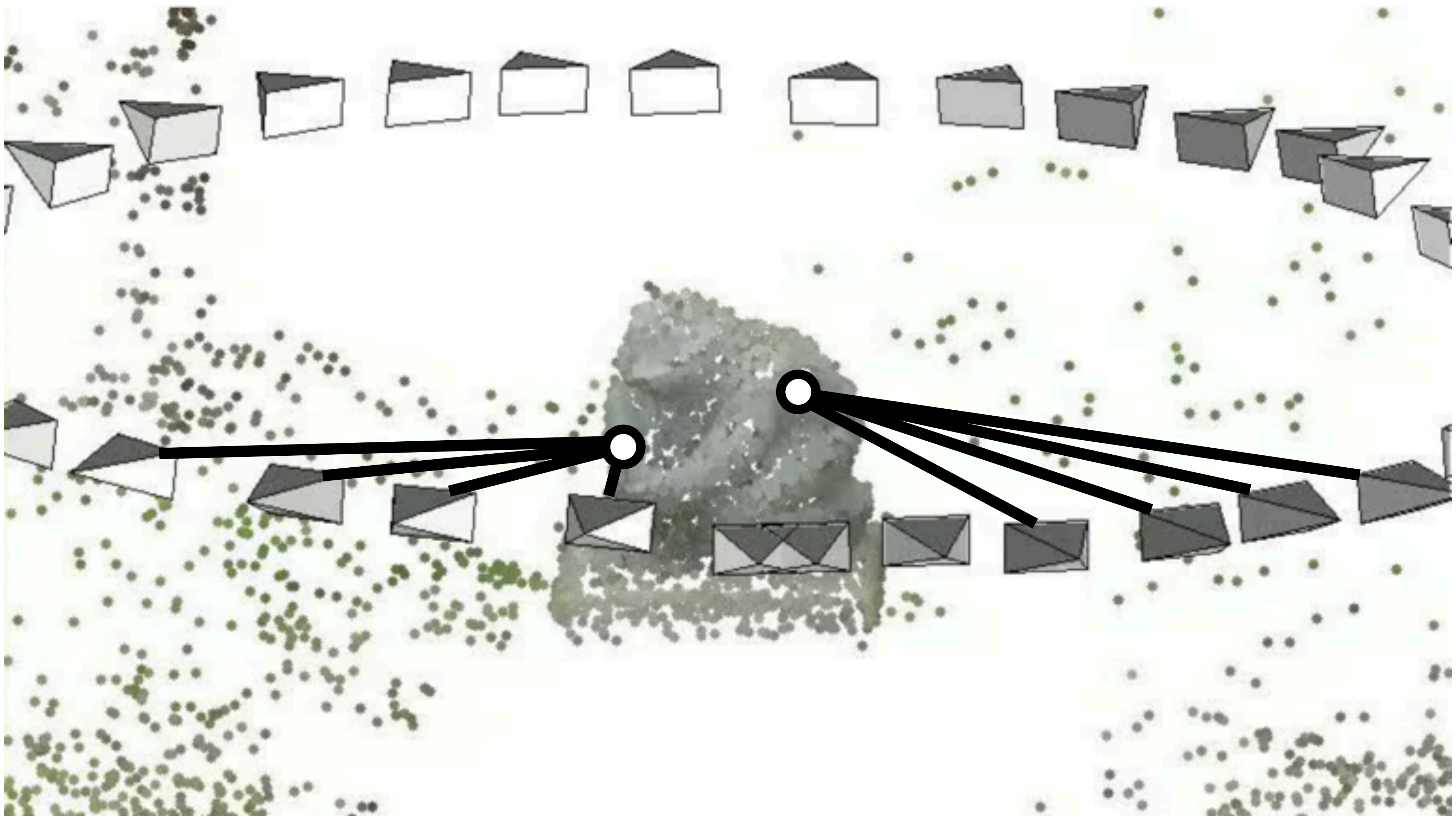
- Add remaining cameras in same way



Structure from Motion

Why “Bundle” Adjustment?

- Can think of bundles of light rays emanating from each 3D point

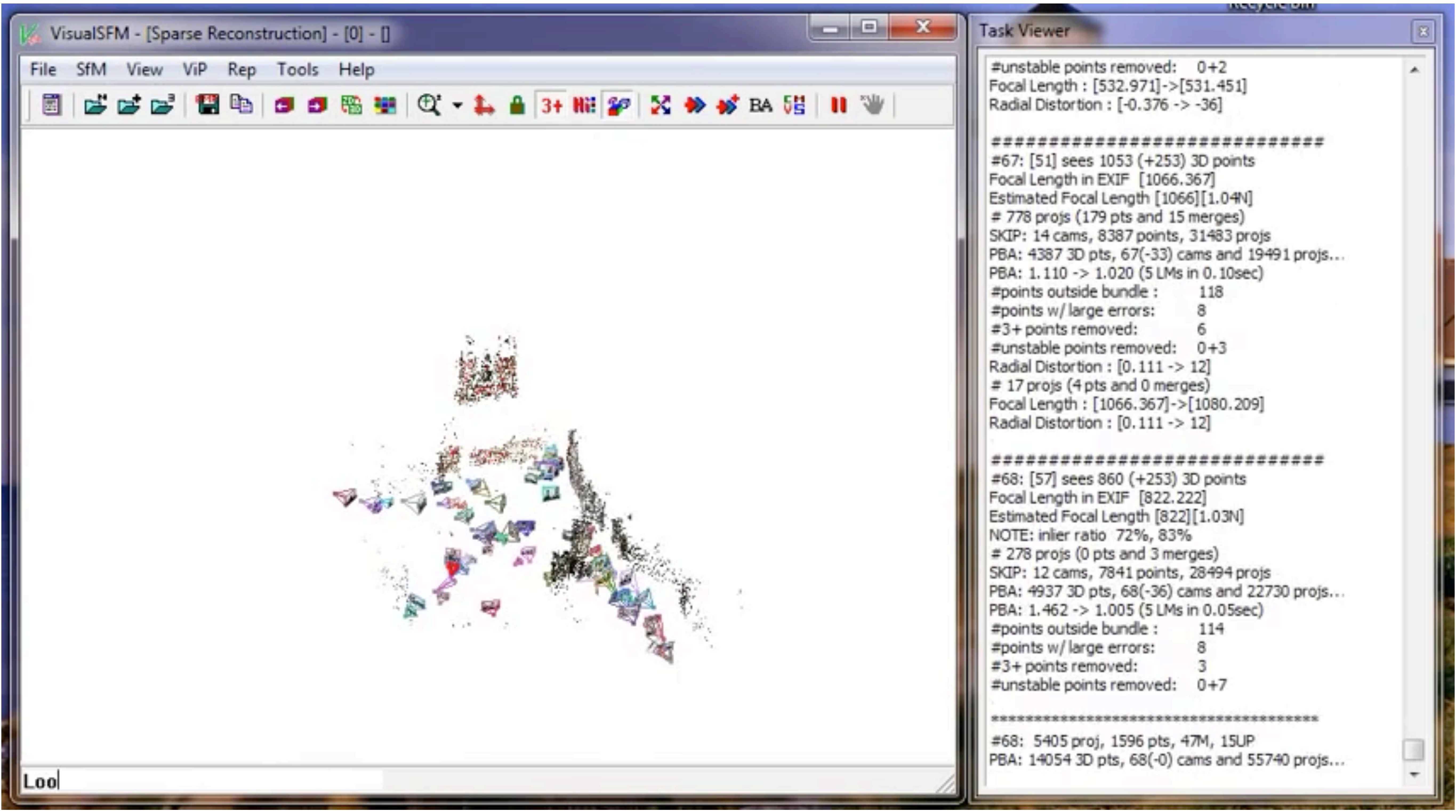


Adjust camera + 3D point positions so that bundles match measured positions (feature points)

SFM recap

- Match features, e.g., SIFT, between all views
- Use RANSAC to reject outliers and estimate F matrices
- Form feature tracks by linking multiview matches
- Select an initialization set, e.g., 3 images with lots of matches and good baseline (parallax)
- Jointly optimize cameras R, t and structure X for this set
- Repeat for each camera:
  - Estimate pose R, t by minimising projection errors with existing X
  - Add 3D points corresponding to the new view and optimize
  - Bundle adjust optimizing over all cameras and structure

Visual SFM



[ccwu.me/vsfm]

Application: 3D from Internet Images

- Reconstruct 3D from unordered photo collections

[Building Rome in a Day, S. Agarwal et al 2009]

Next Lecture

- Dense matching and reconstruction