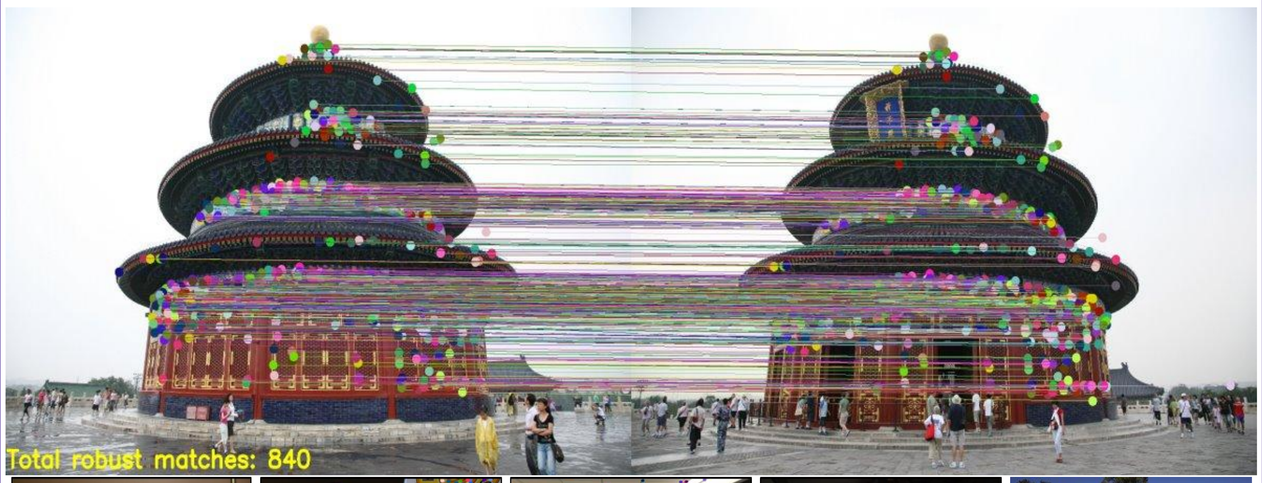
Improving Structure from Motion with Reliable Resectioning

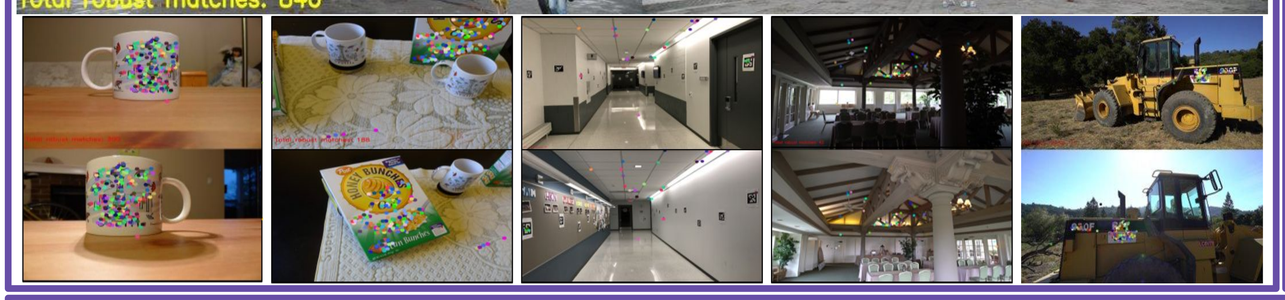
Rajbir Kataria[1], Joseph DeGol[2] and Derek Hoiem[1]
[1]University of Illinois at Urbana-Champaign [2]Microsoft

3DV 2020

Problem

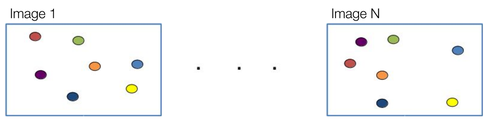
False matches on repeated structures leads to catastrophic failure



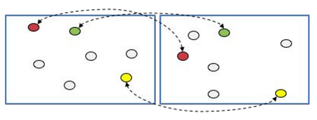


Structure from Motion: Background

(1) Compute and match features: SIFT matching

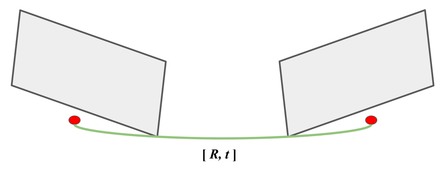


(2) Fit fundamental matrix and match images

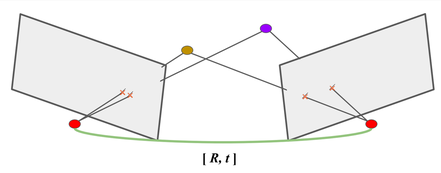


Matching Criteria
$F_{inliers} > \tau$

(3) Initialize reconstruction: Solve for pose between two images

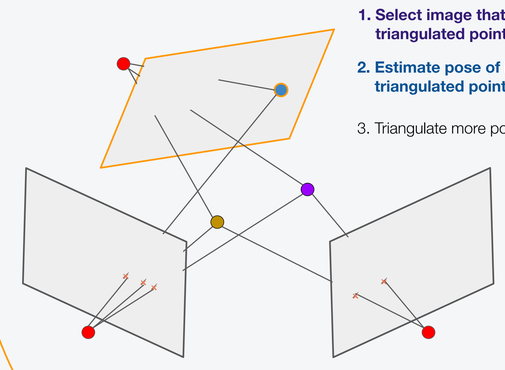


(4) Initialize reconstruction: Triangulate 3D points

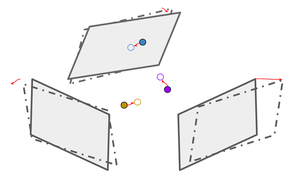


(5) Resectioning: Add additional images and points

1. Select image that views the most triangulated points
2. Estimate pose of image using all the triangulated points (PnP algorithm)
3. Triangulate more points using the added image

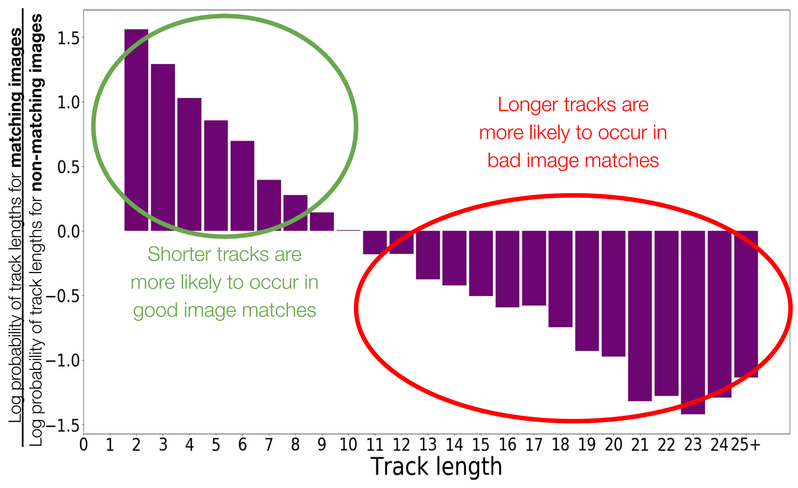


(6) Bundle adjustment



Our method focuses on improving the resectioning process

Method



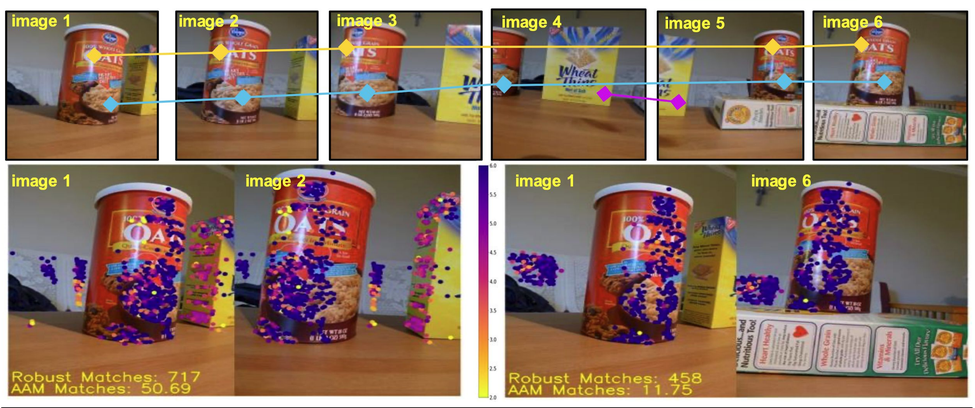
Hyperparameter: 0.5

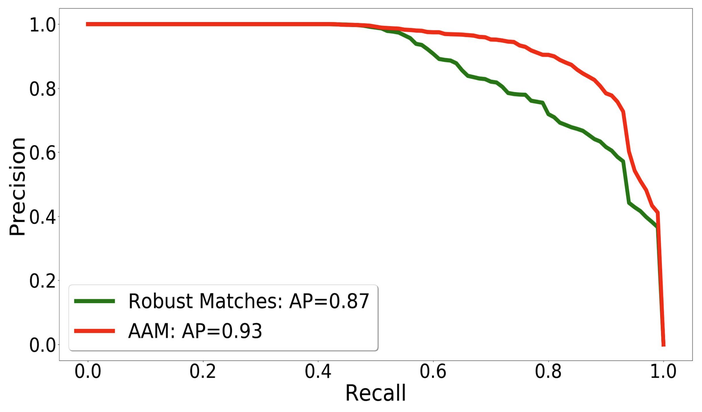
Track length of k^{th} track

Number of AAM matches between images i and j

$$S_{AAM}(i,j) = \sum_{t_k \in T_{ij}} \gamma^{|t_k|-2}$$

Common tracks between images i and j

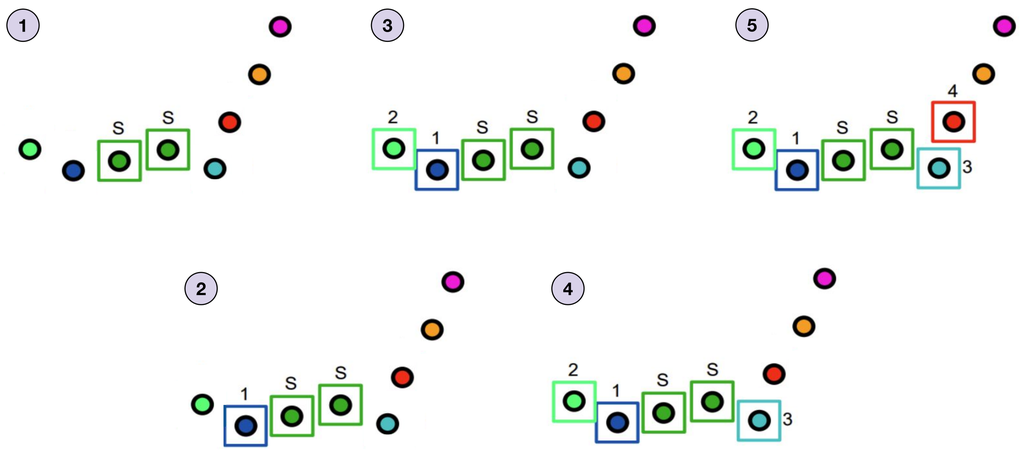


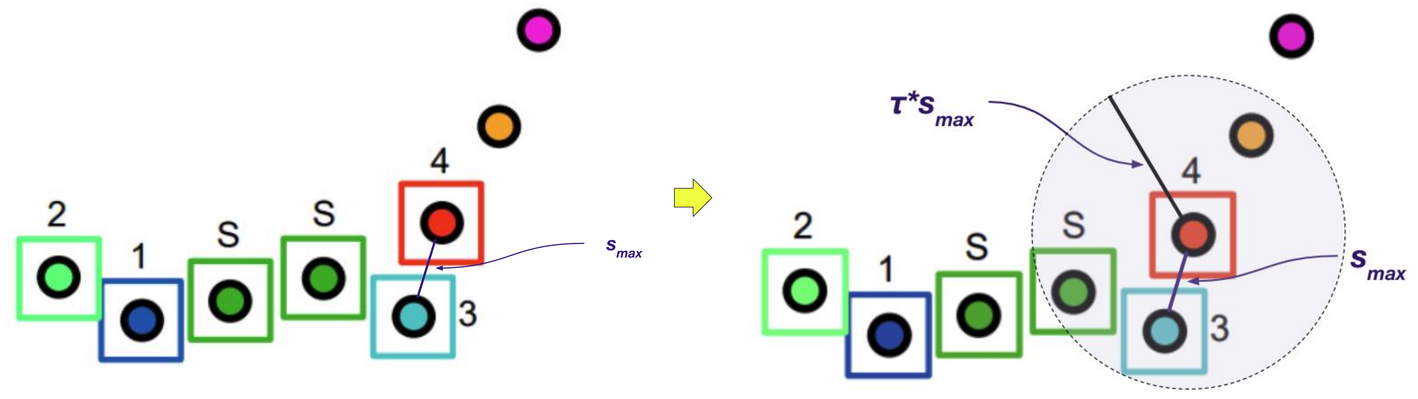


(1) We observe that shorter tracks are more likely in good image matches for scenes with duplicate structures.

(2) We give longer tracks less importance by discounting all matches based on their track length. This is analogous to how tf-idf downweights common features in retrieval.

(3) Our similarity measure, ambiguity-adjusted matches (AAM), outperforms robust matches in discriminating duplicate structures.





Local resectioning order: Add the most similar image to any reconstructed image where similarity is measured in AAM.

Local pose estimation: Determine the similarity value between the resectioned image and it's most similar reconstructed image (s_{max}). s_{max} and τ (pre-set parameter) determine the set of reliable reconstructed images. Points viewed by these images are used for pose estimation.

We use points from a smaller set of reliable images to determine resectioning order and pose estimation

Results

Our method improves standard pipelines

	OpenSfM	OpenSfM w/ Our Resectioning	COLMAP	COLMAP w/ Our Resectioning
Duplicate Structures Dataset	6 Failures	6 Successes	6 Failures	2 Failures, 4 Successes
UIUCTag Dataset	9 Failures, 3 Partial Successes, 4 Successes	3 Failures, 13 Successes	7 Failures, 2 Partial Successes, 7 Successes	5 Failures, 3 Partial Successes, 8 Successes
TanksAndTemples Dataset	4 Failures, 3 Successes	1 Failure, 6 Successes	1 Failure, 6 Successes	7 Successes

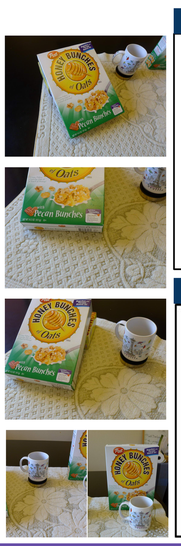
Our method outperforms other disambiguation methods

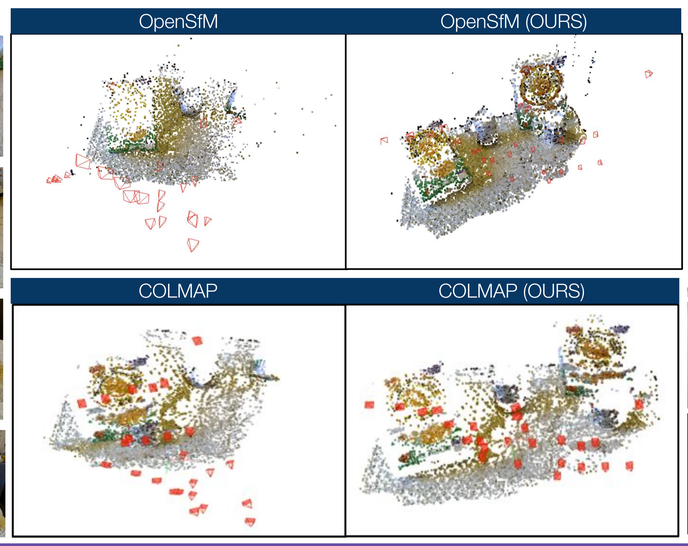
	OpenSfM w/ Our Resectioning	Yan et al.	Wilson et al.
Duplicate Structures Dataset	6 Successes	1 Failure, 2 Partial Successes, 3 Successes	6 Failures
UIUCTag Dataset	3 Failures, 13 Successes	6 Failures, 9 Partial Successes, 1 Success	10 Failures, 5 Partial Successes, 1 Success
TanksAndTemples Dataset	1 Failure, 6 Successes	2 Failures, 5 Successes	4 Failures, 3 Successes

Local pose estimation has the largest positive impact

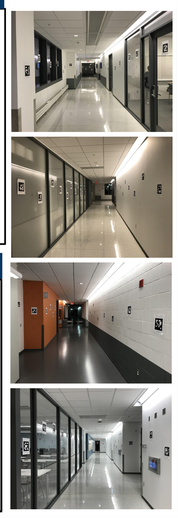
	OpenSfM w/ Our Resectioning	w/o AAM	w/o Local resectioning order	w/o Local pose estimation
Duplicate Structures Dataset	6 Successes	3 Failures, 3 Successes	3 Failures, 3 Successes	6 Failures
UIUCTag Dataset	3 Failures, 13 Successes	1 Failure, 15 Successes	4 Failures, 12 Successes	10 Failures, 3 Partial Successes, 3 Successes
TanksAndTemples Dataset	1 Failure, 6 Successes	1 Failure, 6 Successes	7 Successes	4 Failures, 3 Successes

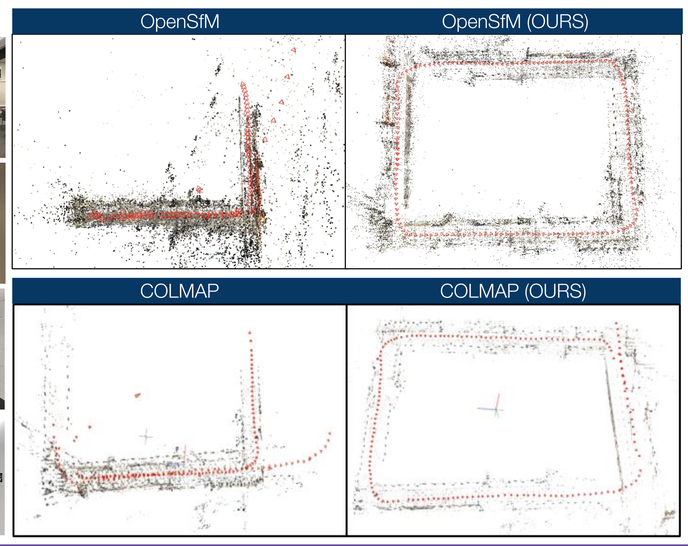
Cereal (Duplicate Structures)





ece_floor3_loop_cw (UIUCTag)





Courthouse (TanksAndTemples)