

Towards Vision Based Robots for Monitoring Built Environments

Joseph DeGol

Chair: Dr. Derek Hoiem

Committee: Dr. Timothy Bretl

Dr. Mani Golparvar-Fard

Dr. David Forsyth

Dr. Sudipta Sinha, MSR

May 8, 2018

PhD Defense

University of Illinois Urbana-Champaign

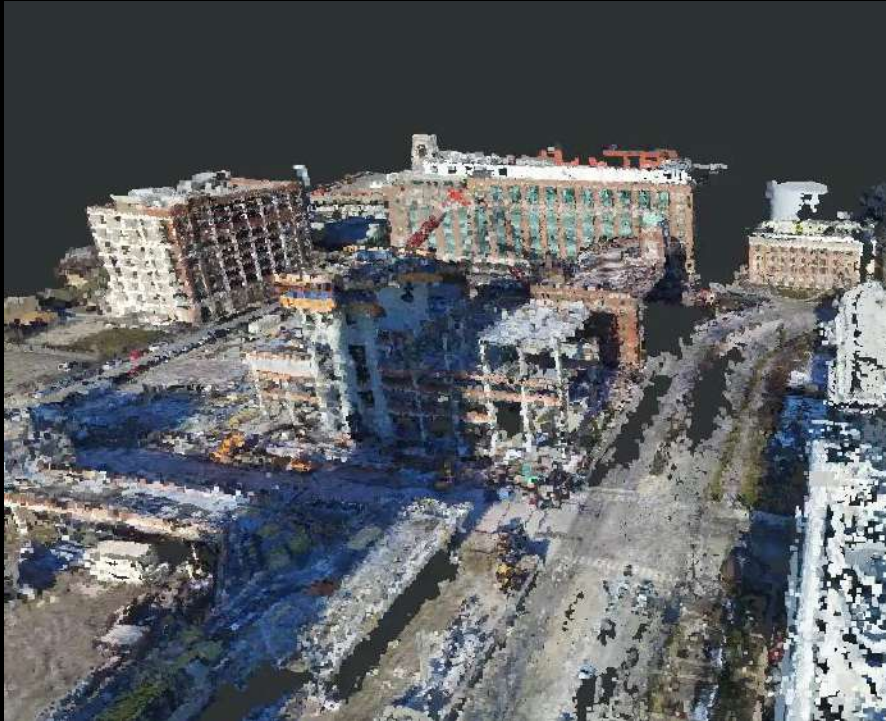
Vision Based Robots Monitoring Built Environments



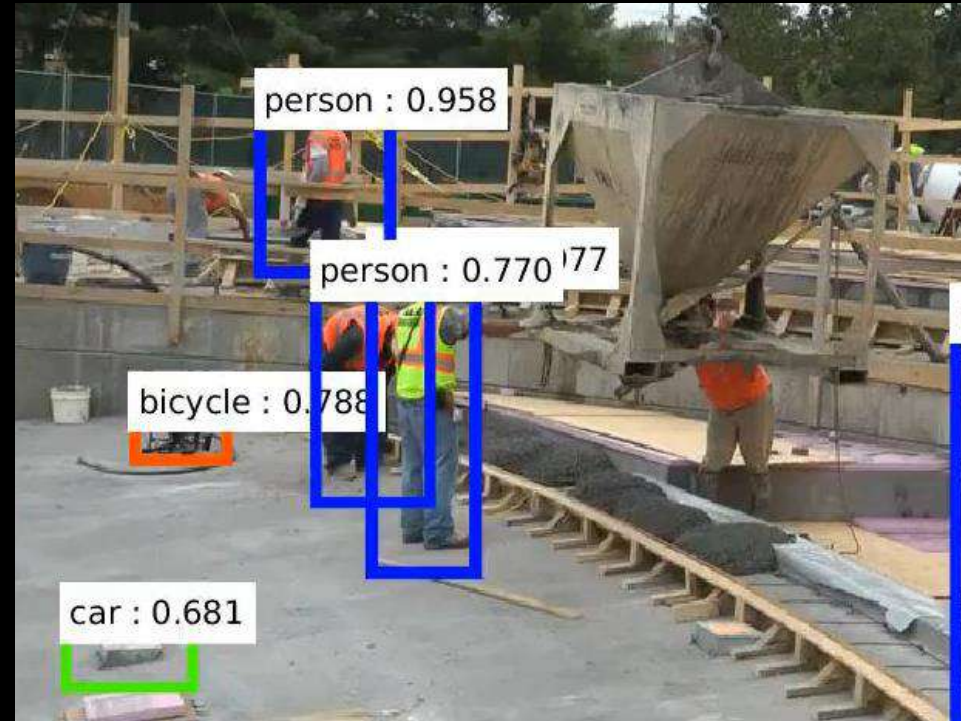
Vision Based Robots Monitoring Built Environments



Vision Based Robots Monitoring Built Environments

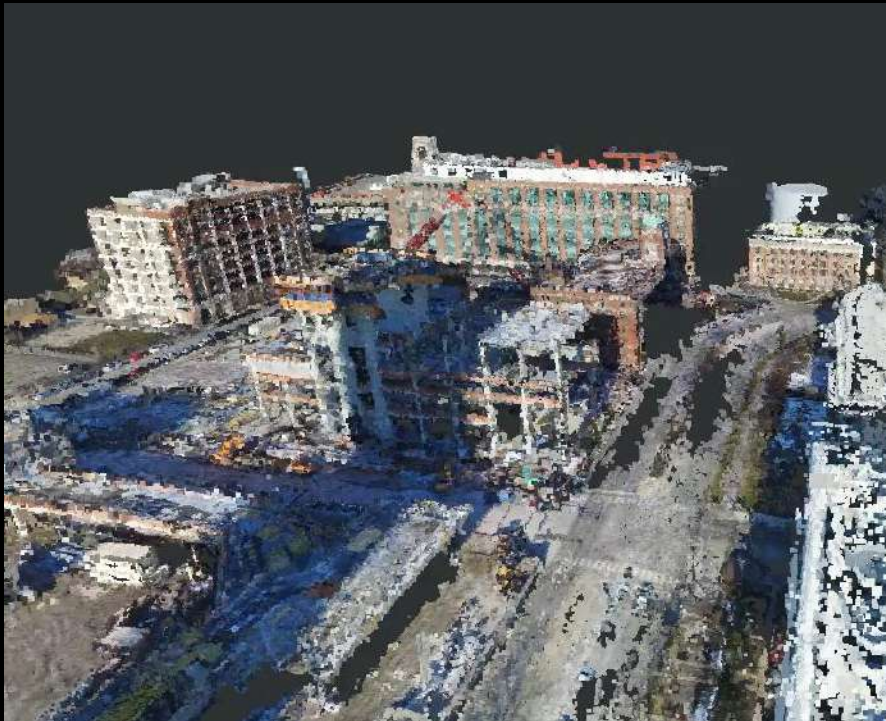


3D Reconstruction to Monitor Progress



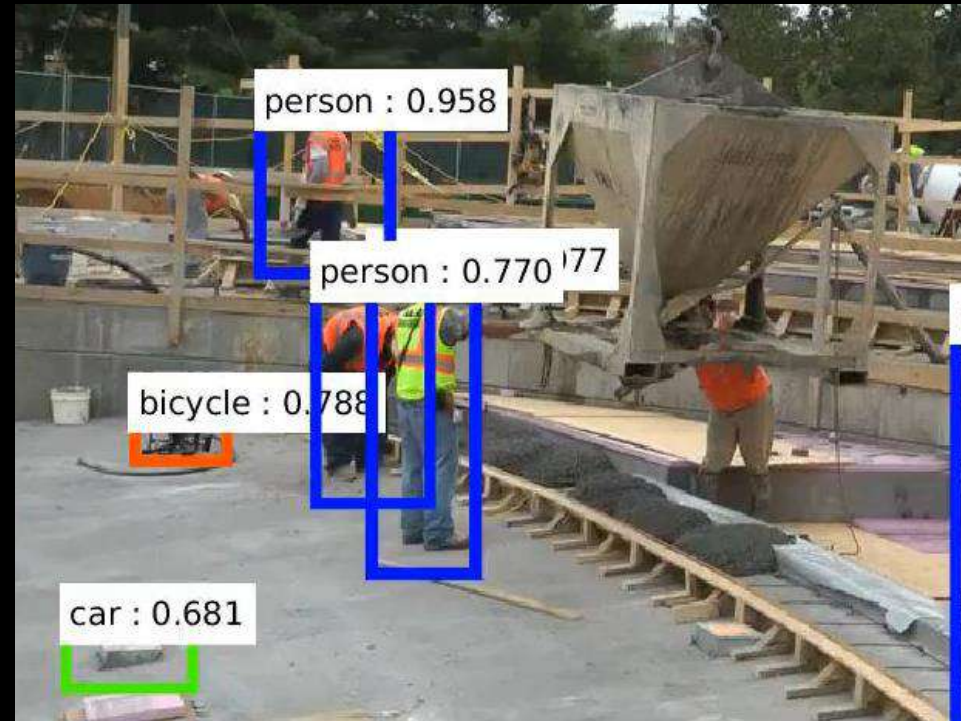
Worker, Equipment, and Material Tracking

Vision Based Robots Monitoring Built Environments



Large projects average 20 month delays
and 80% cost overruns

McKinsey & Company (2015)



991 Workers Lost Lives in 2016

[osha.gov/oshstats/commonstats.html](https://www.osha.gov/oshstats/commonstats.html)

Vision Based Robots Monitoring Built Environments

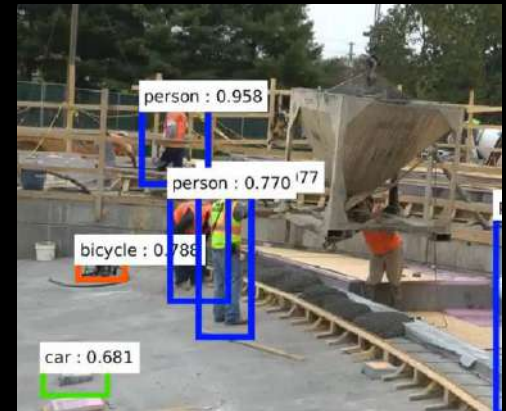
Data Capture



Mapping



Analysis



Vision Based Robots Monitoring Built Environments

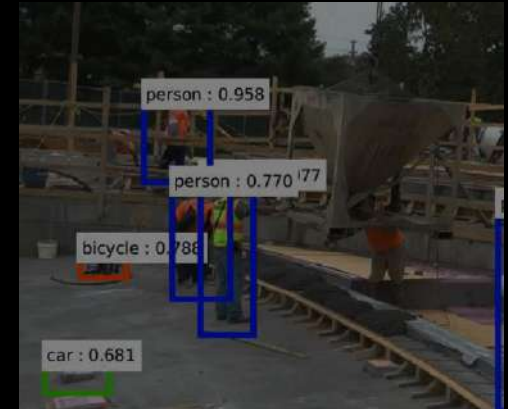
Data Capture



Mapping



Analysis



Structure from Motion

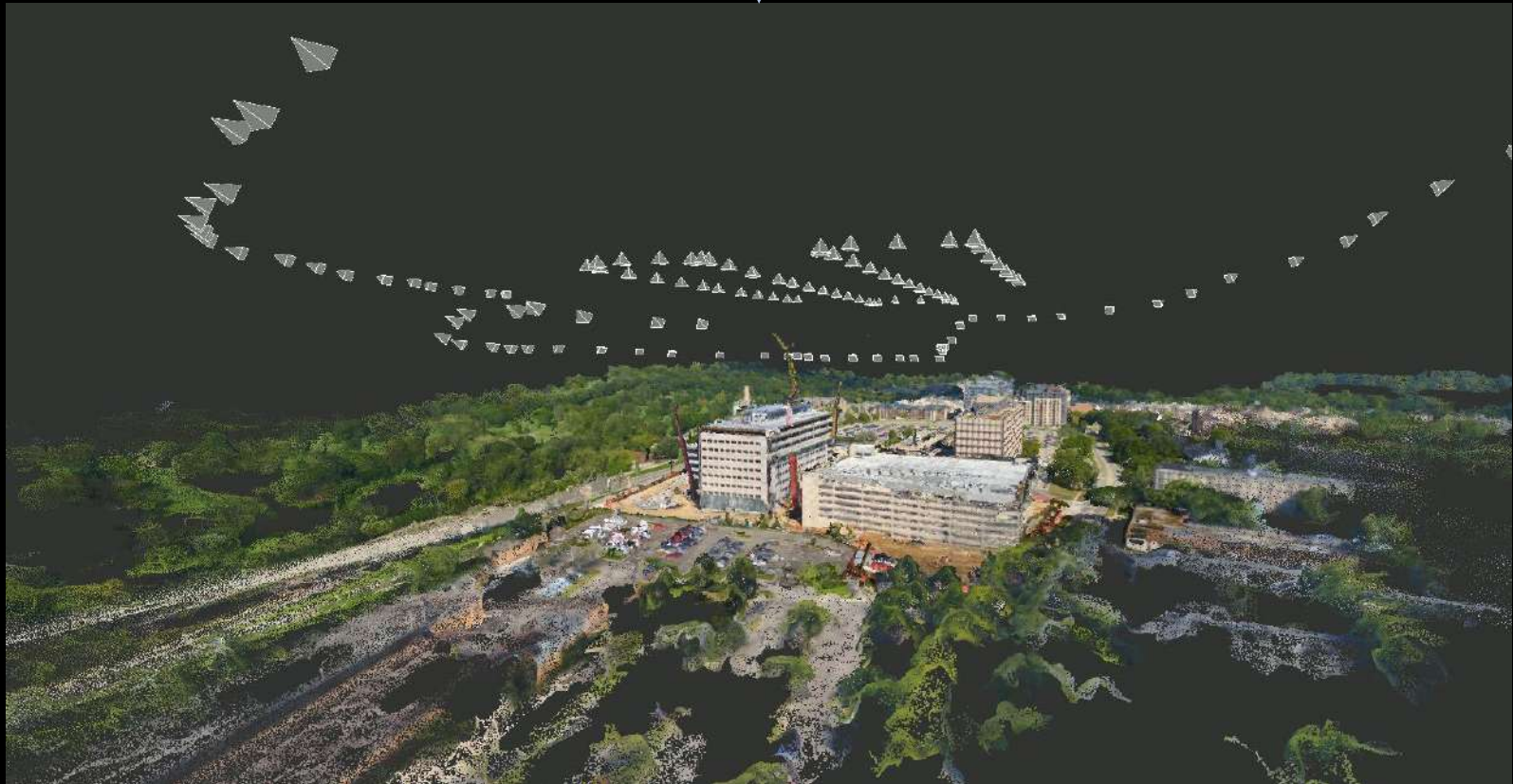
Structure from Motion: From Image Collections

Image Collection



Structure
from Motion

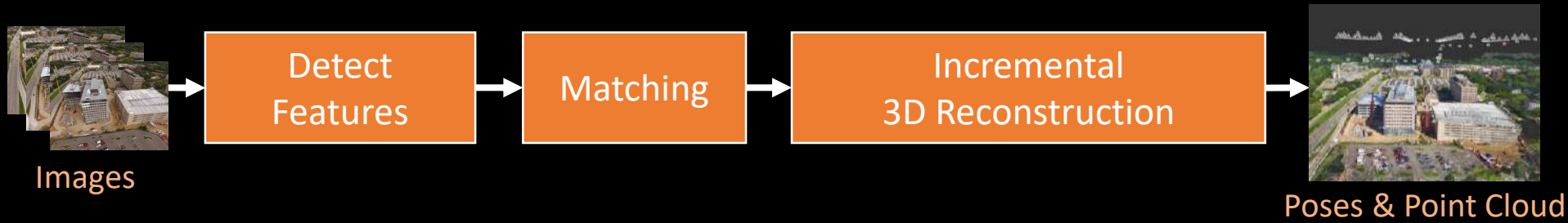
Structure from Motion: To 3D Reconstructions



3D Reconstruction

(Image Locations and 3D Points)

Structure from Motion: And Everything in Between



Structure from Motion: And Everything in Between

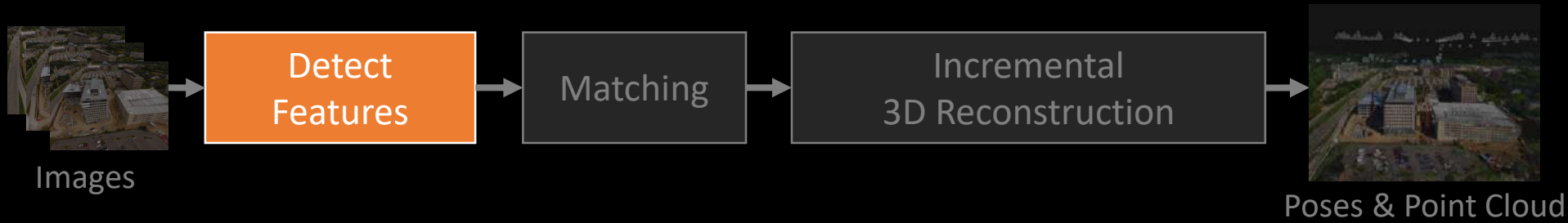
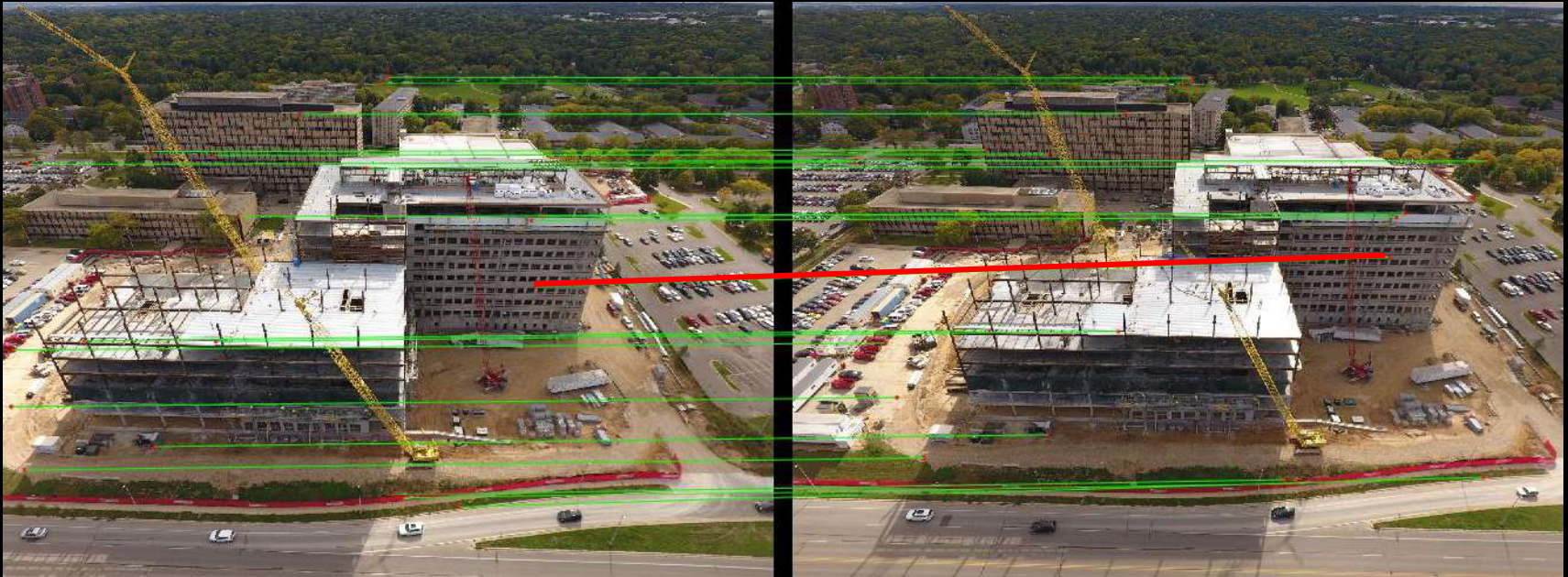
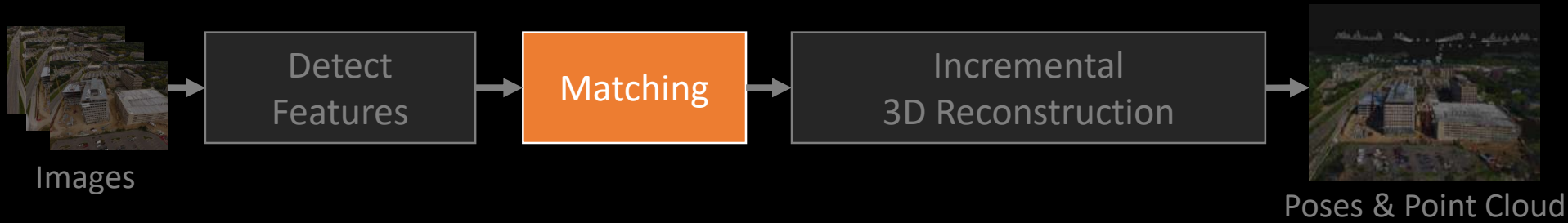


Image Locations with Texture

Structure from Motion: And Everything in Between



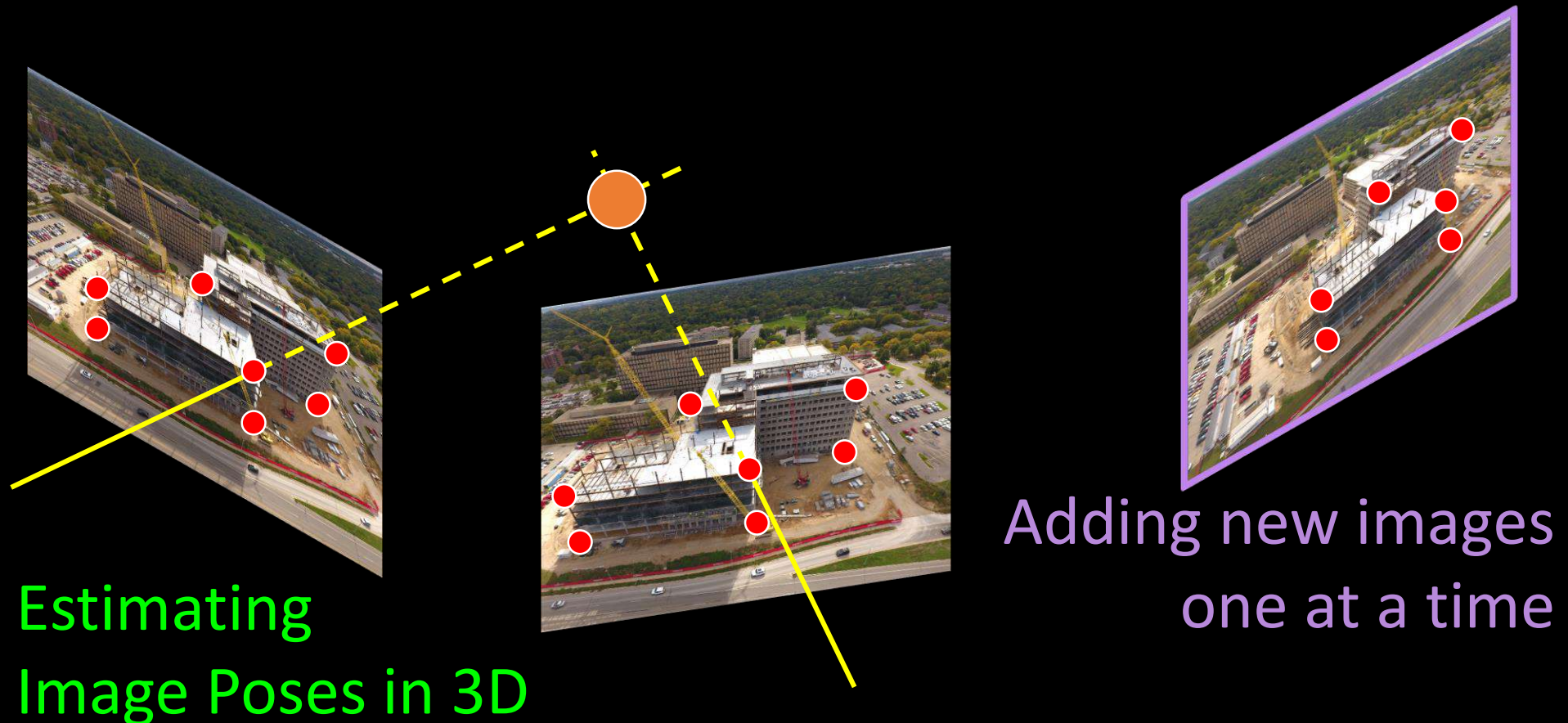
The SAME location across images

Bad matches happen

Structure from Motion: And Everything in Between



Triangulating 3D Point Positions

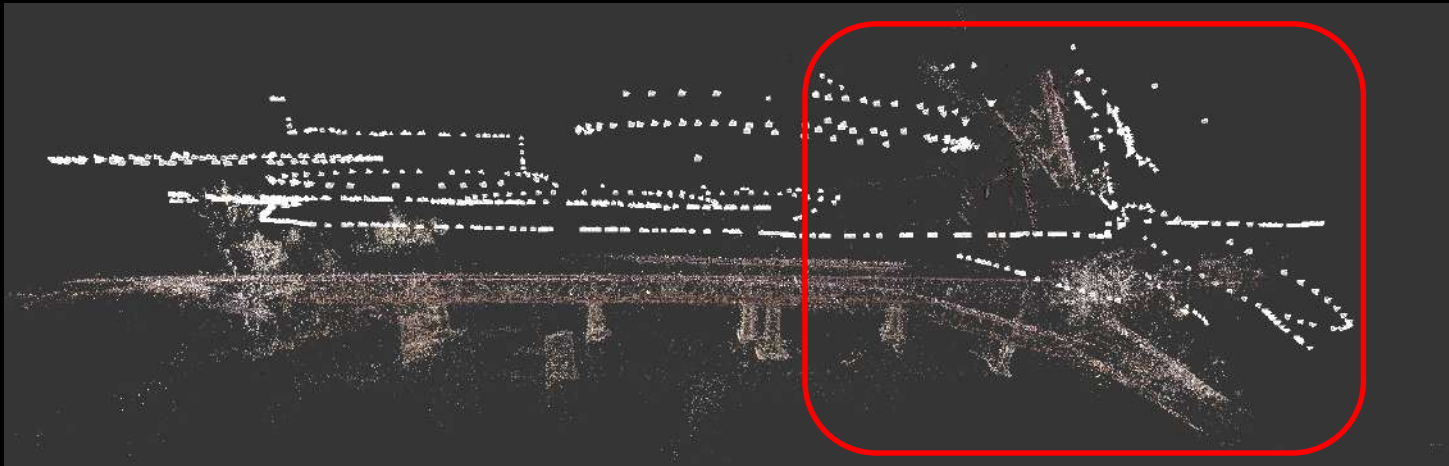


Sometimes Structure from Motion Fails

Top View



Side View



Vision Based Robots Monitoring Built Environments

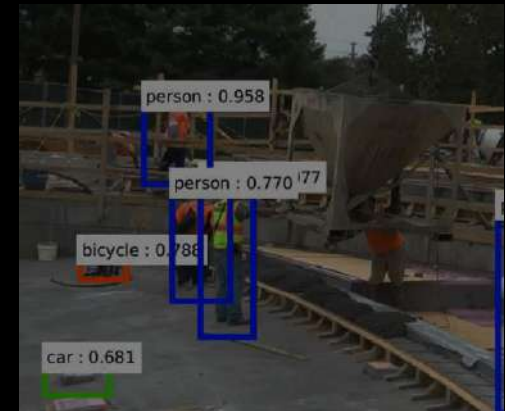
Data Capture



Mapping



Analysis

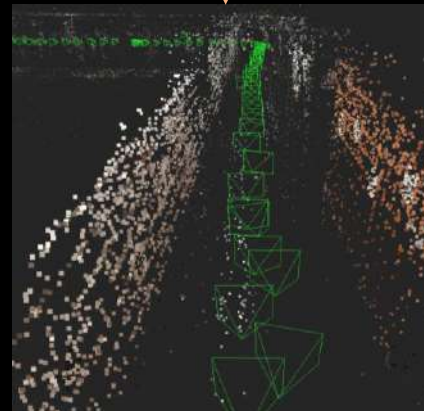


Structure from Motion

Overcoming
Failures



Simulating Features to
Predict Mapping Failure



Using Markers for
Improved 3D Mapping

Vision Based Robots Monitoring Built Environments

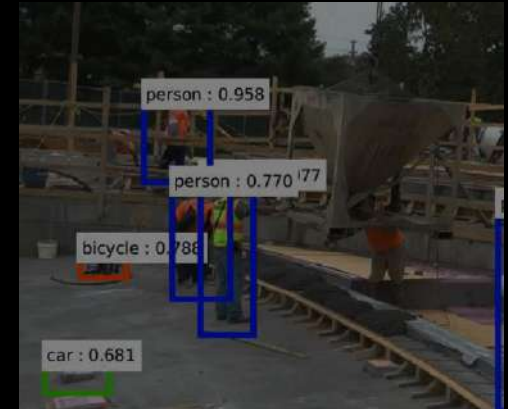
Data Capture



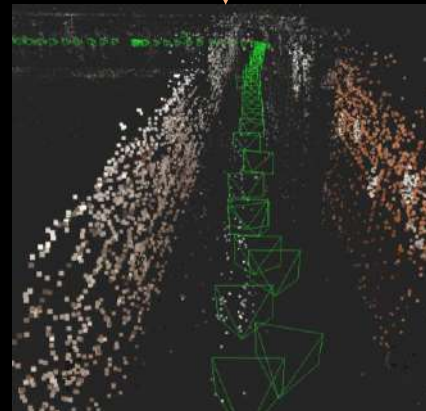
Mapping



Analysis



Structure from Motion



Using Markers for
Improved 3D Mapping

Improved Structure from Motion Using Fiducial Marker Matching

*Joseph DeGol,
Timothy Bretl,
Derek Hoiem*

Introduction

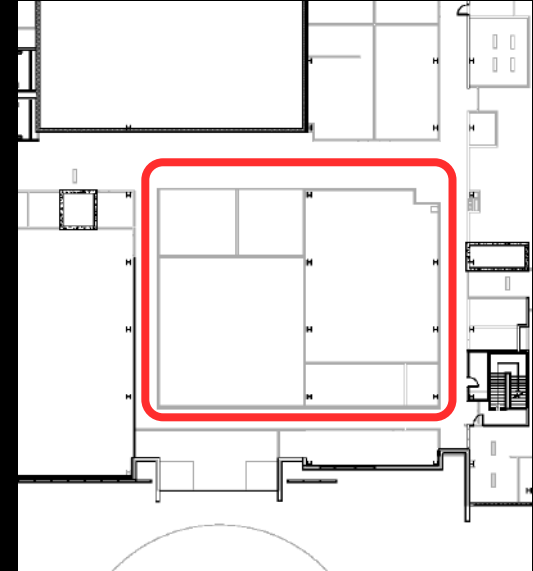
Submitted to ECCV 2018

Another Structure from Motion Example

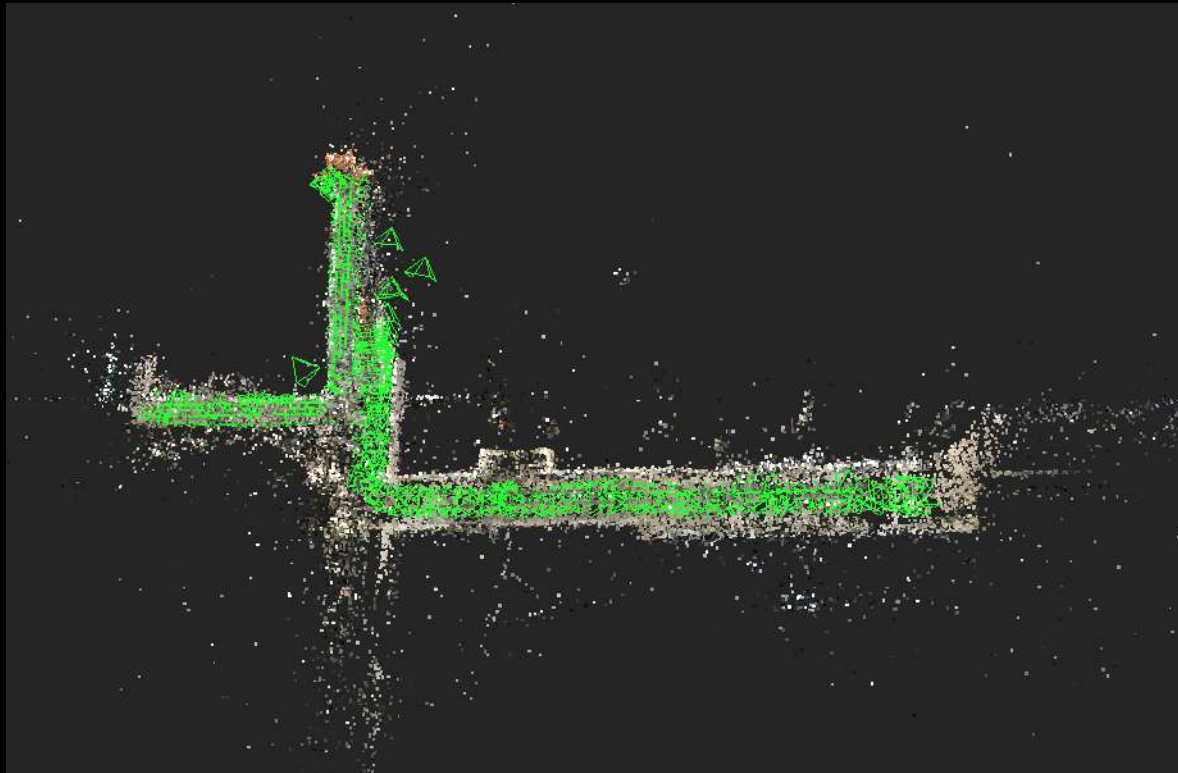
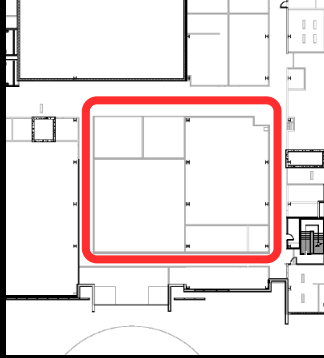
Image Collection



Floor Plan



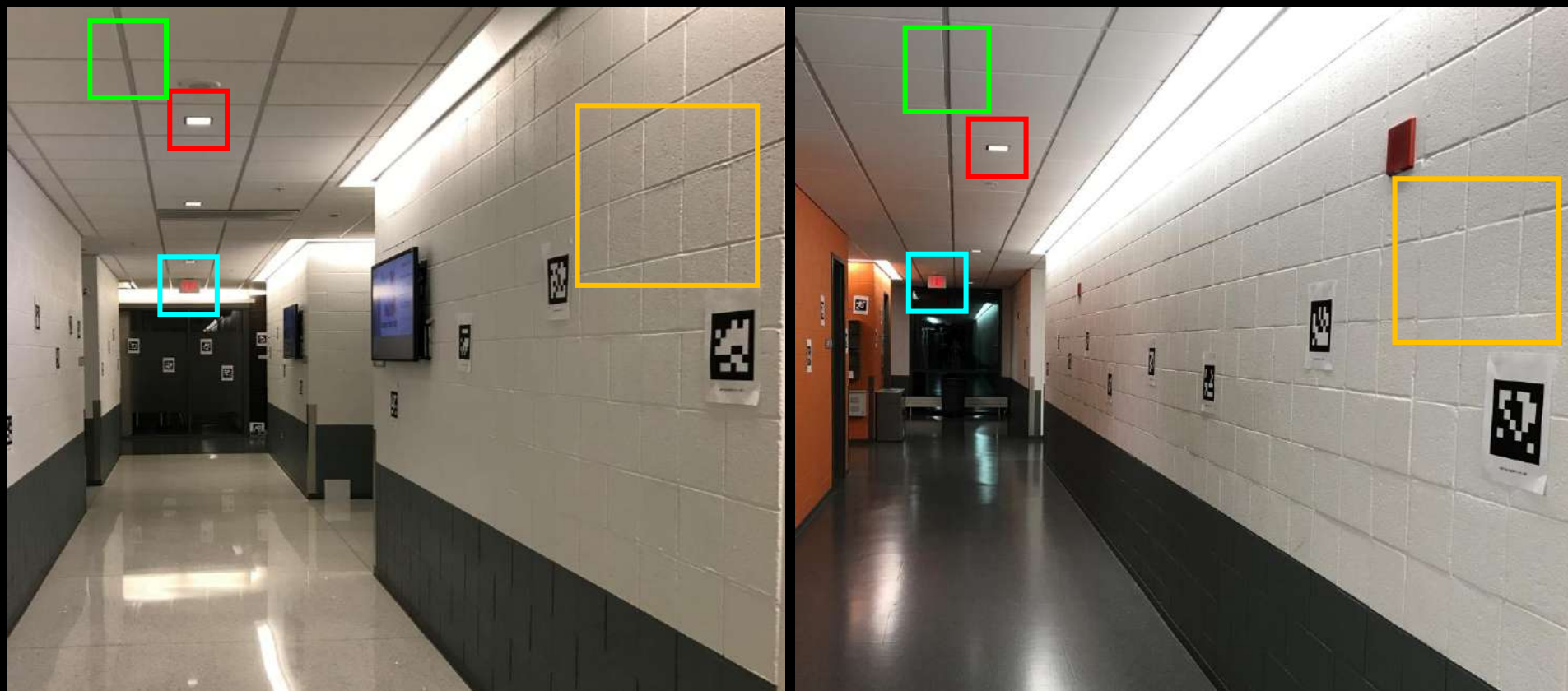
Structure from Motion Failure



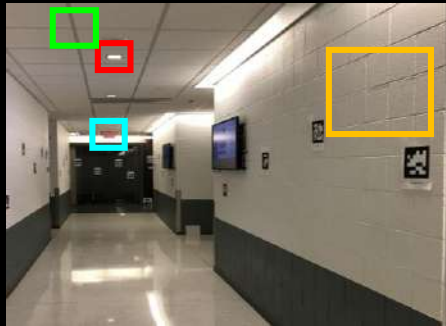
Failed 3D Reconstruction

Reasons for Structure from Motion Failures

Repetitive Surfaces: *repetitive features are confused*



Reasons for Structure from Motion Failures



Reflective Surfaces: *reflected feature motion inconsistent with scene motion*



Reasons for Structure from Motion Failures



Textureless Surfaces: *few features to track*



Reasons for Structure from Motion Failures

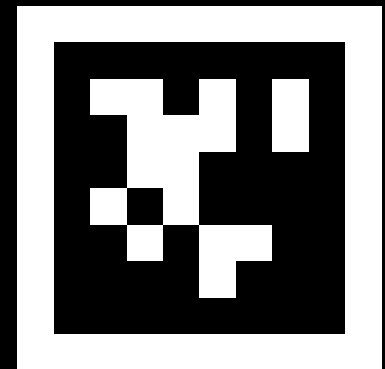


Problem:

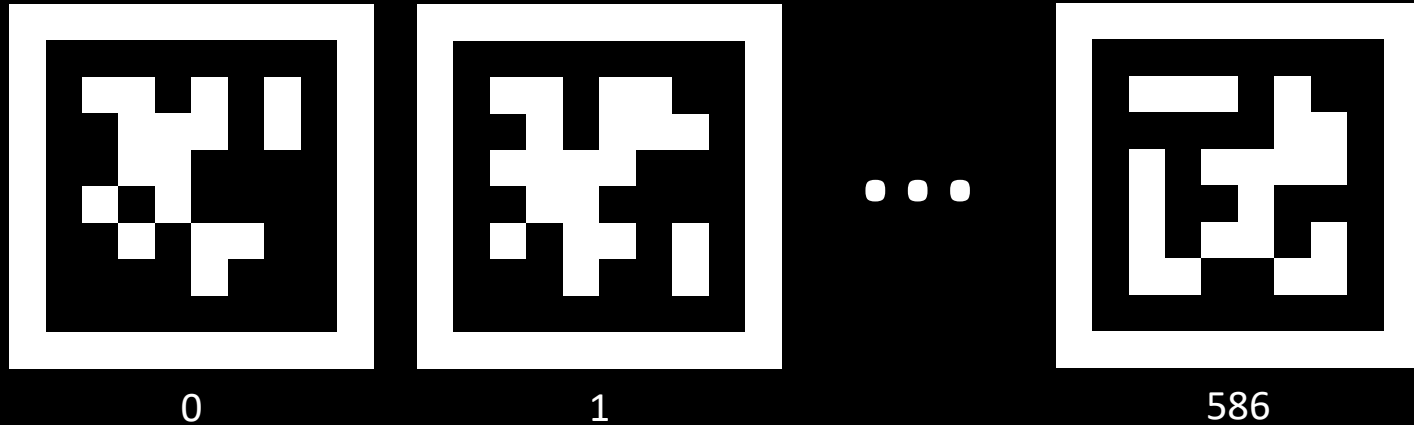
Lack of Stable Features

Solution:

Markers Provide Stable Features

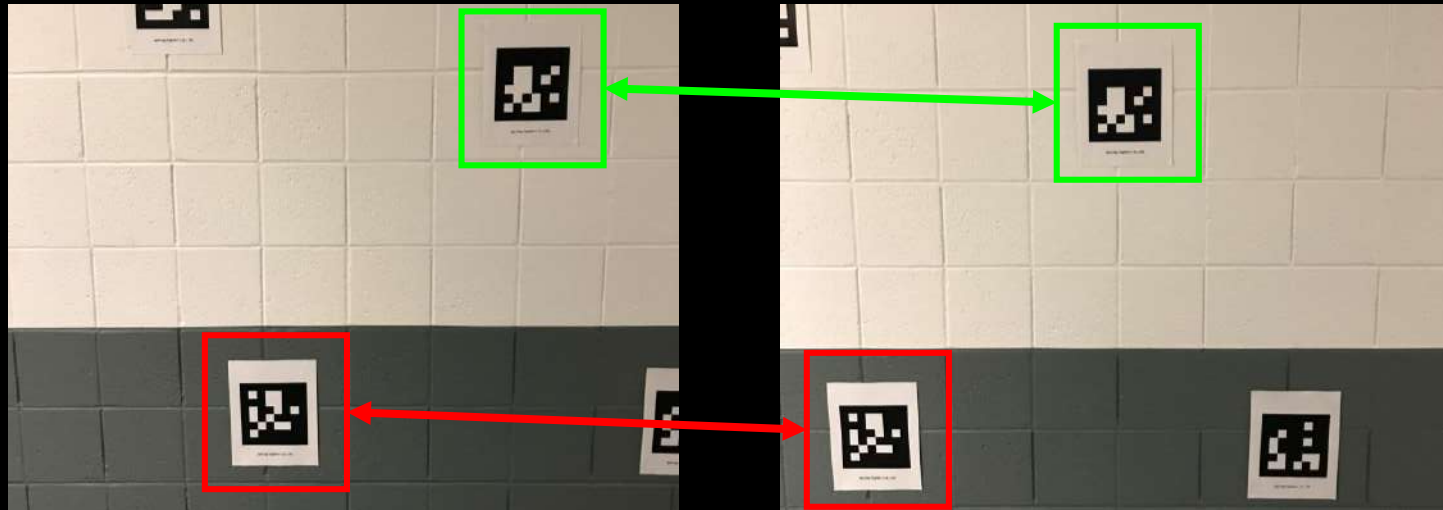


Unique ID

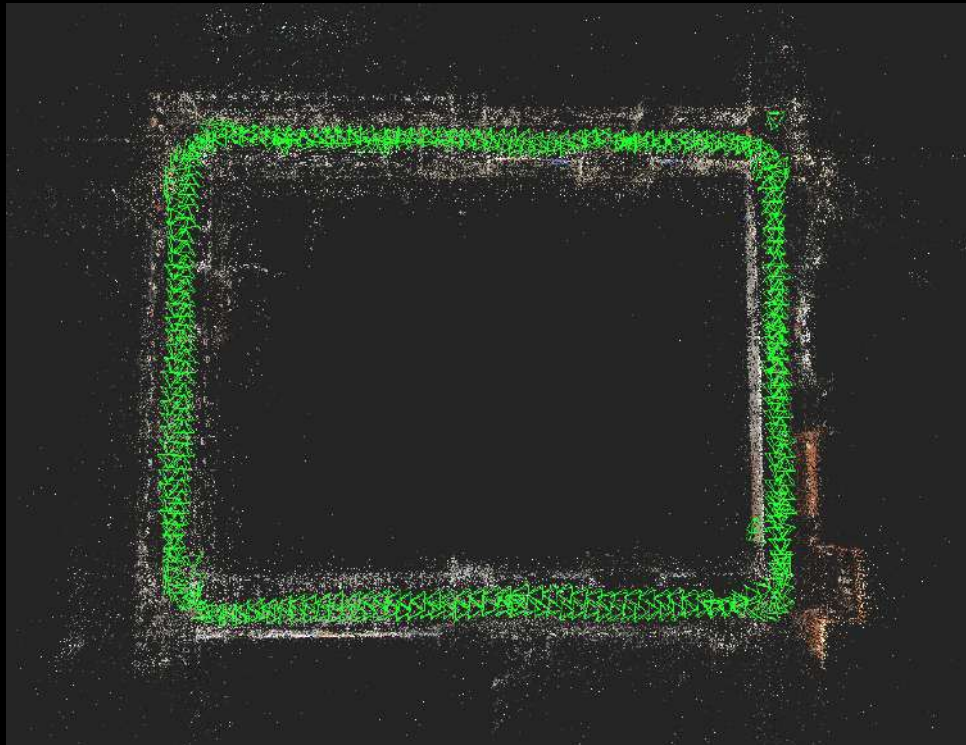
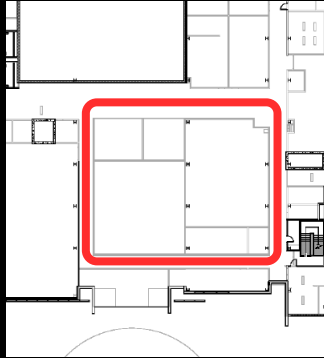
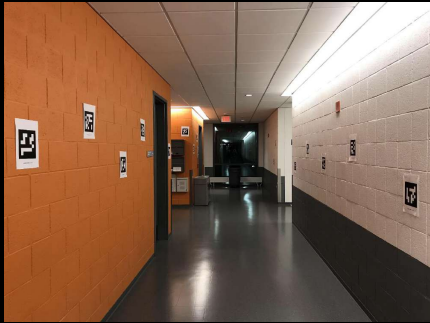


Almost Perfect Matching

False Positive Rate of 0.000044% [Wang and Olson]



Improved Structure from Motion from Markers

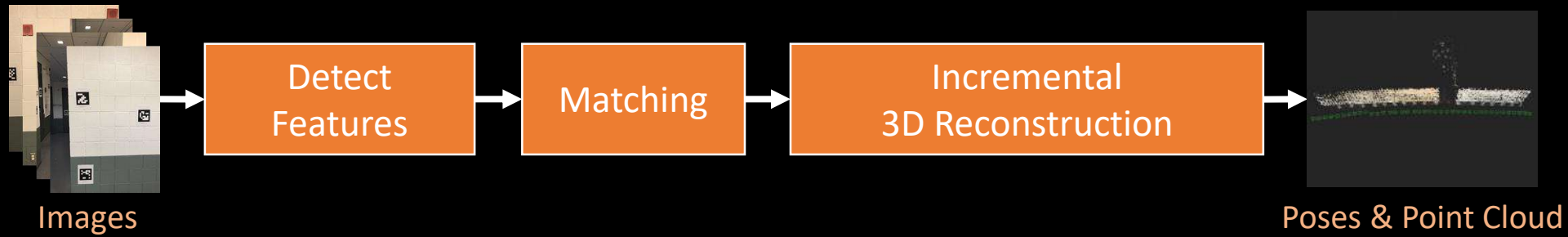


Successful 3D Reconstruction

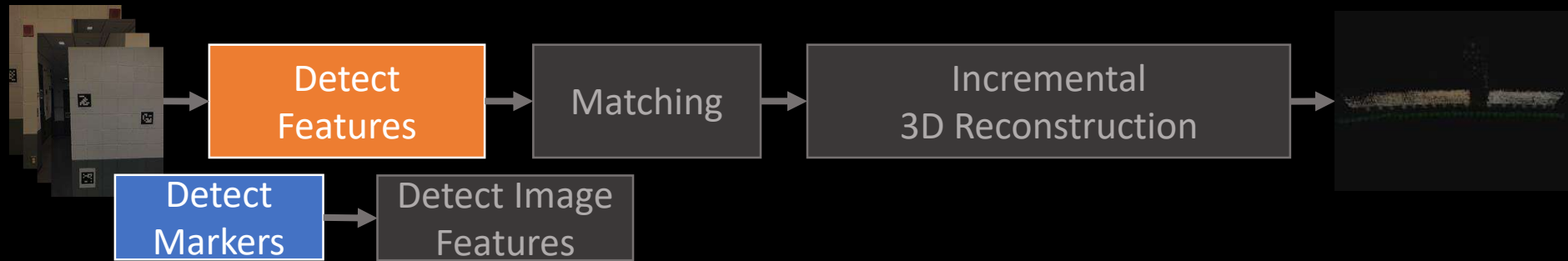
Improved Structure from Motion Using Fiducial Marker Matching

Using Marker Matches

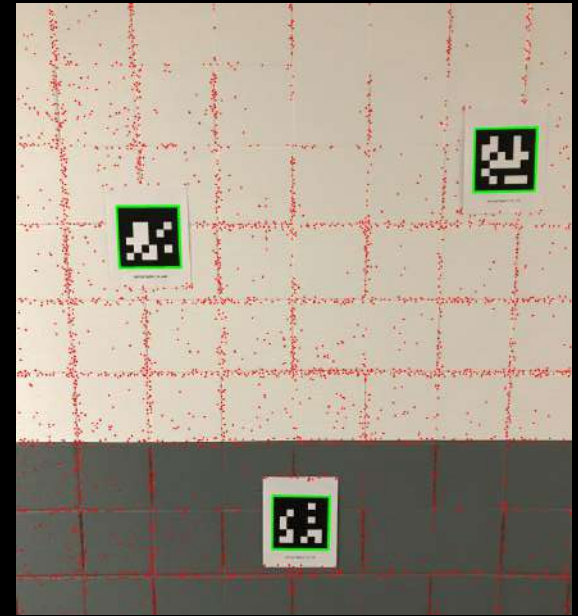
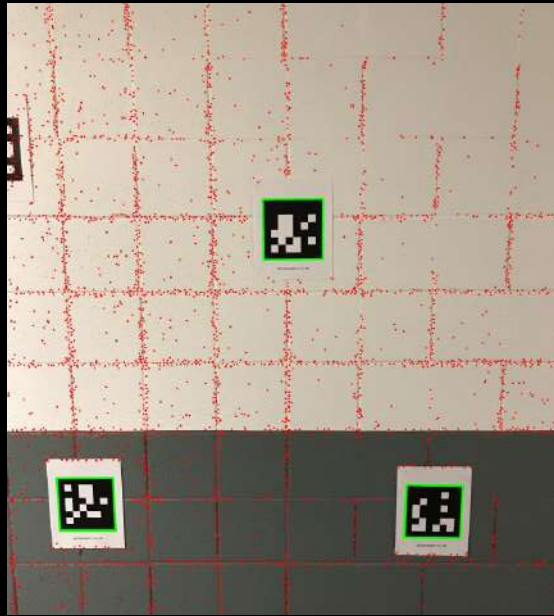
Structure from Motion Pipeline



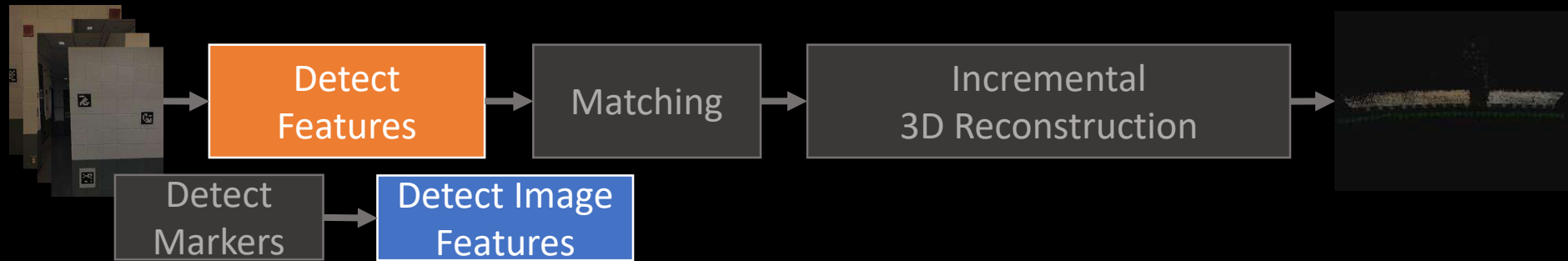
Structure from Motion Pipeline



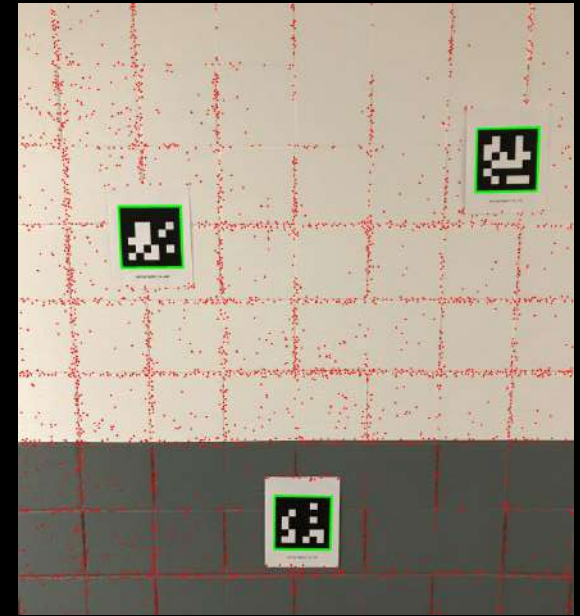
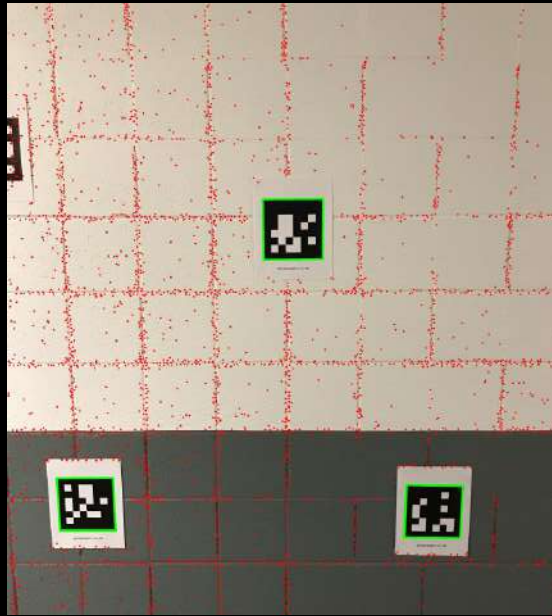
Detect Markers



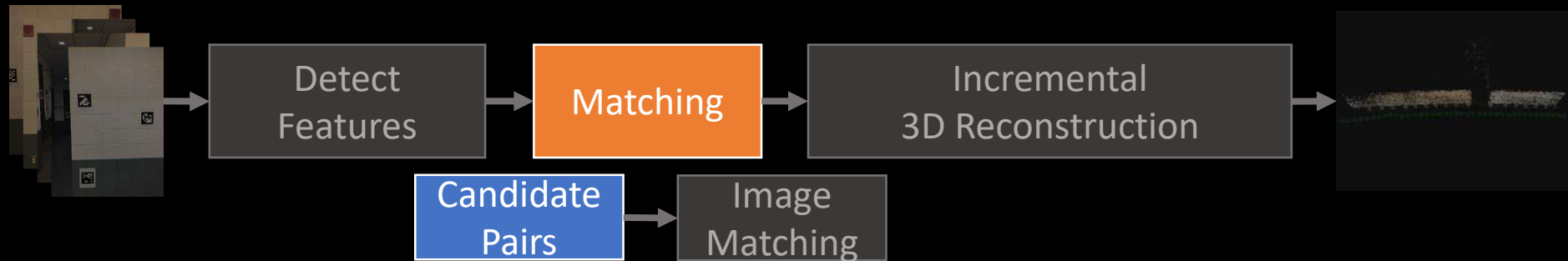
Structure from Motion Pipeline



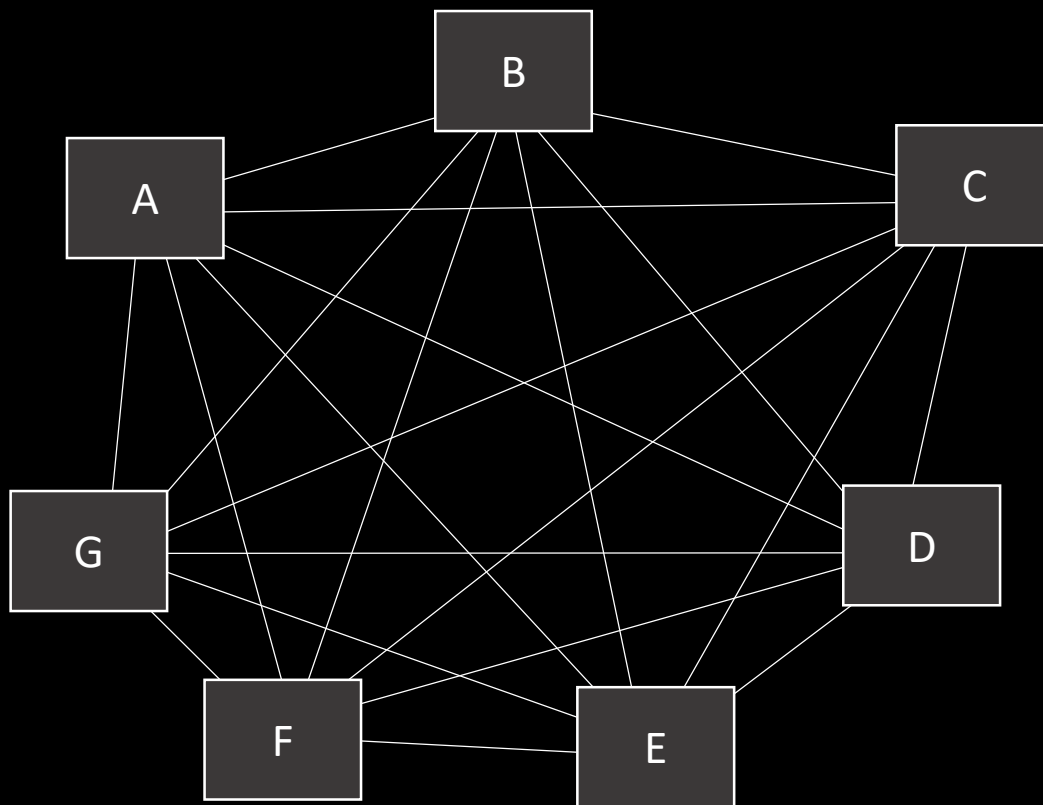
Detect Image Features



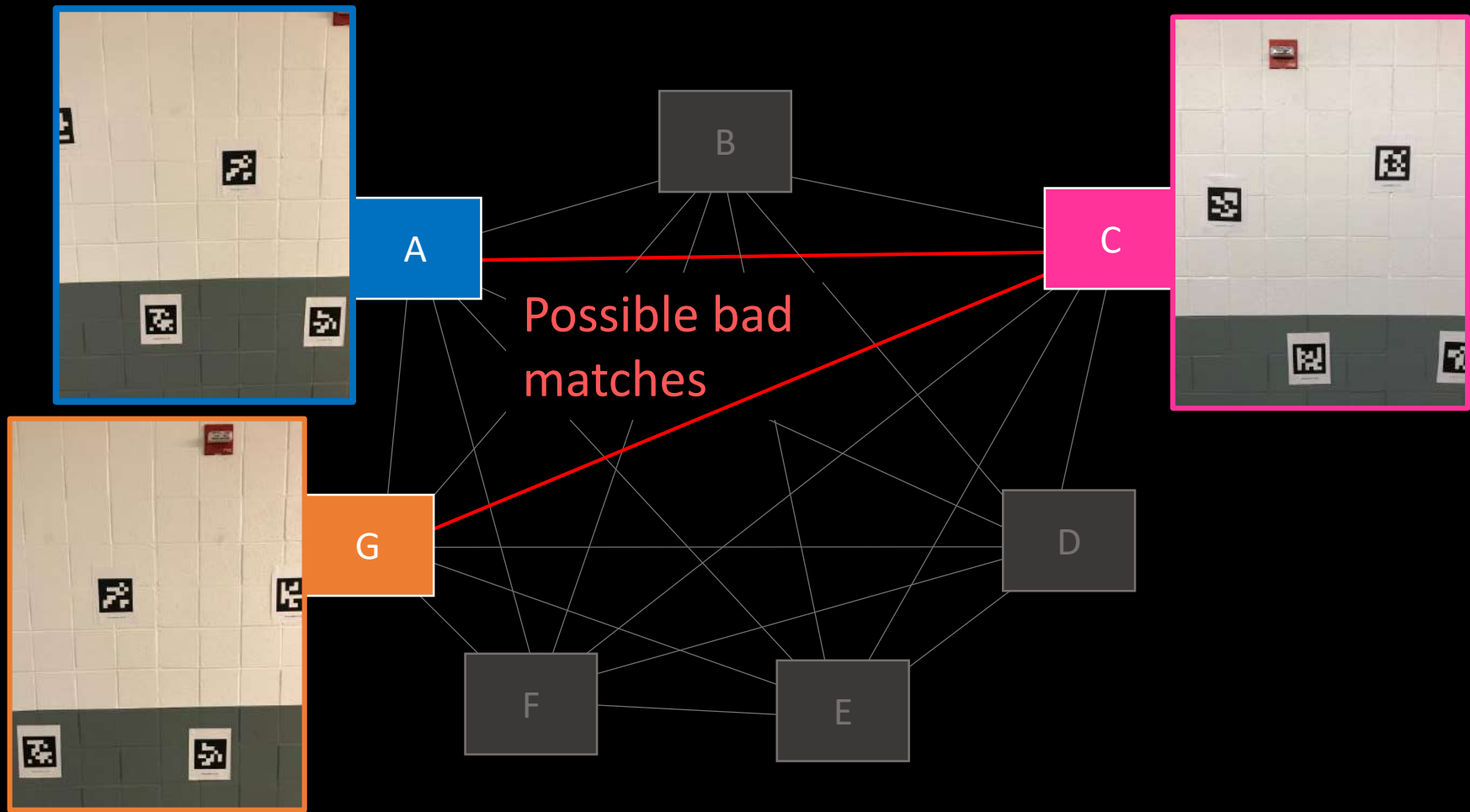
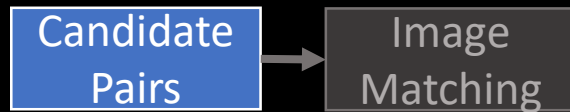
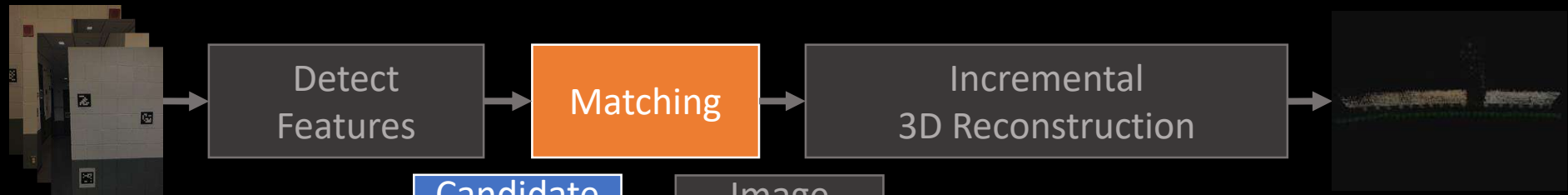
Structure from Motion Pipeline



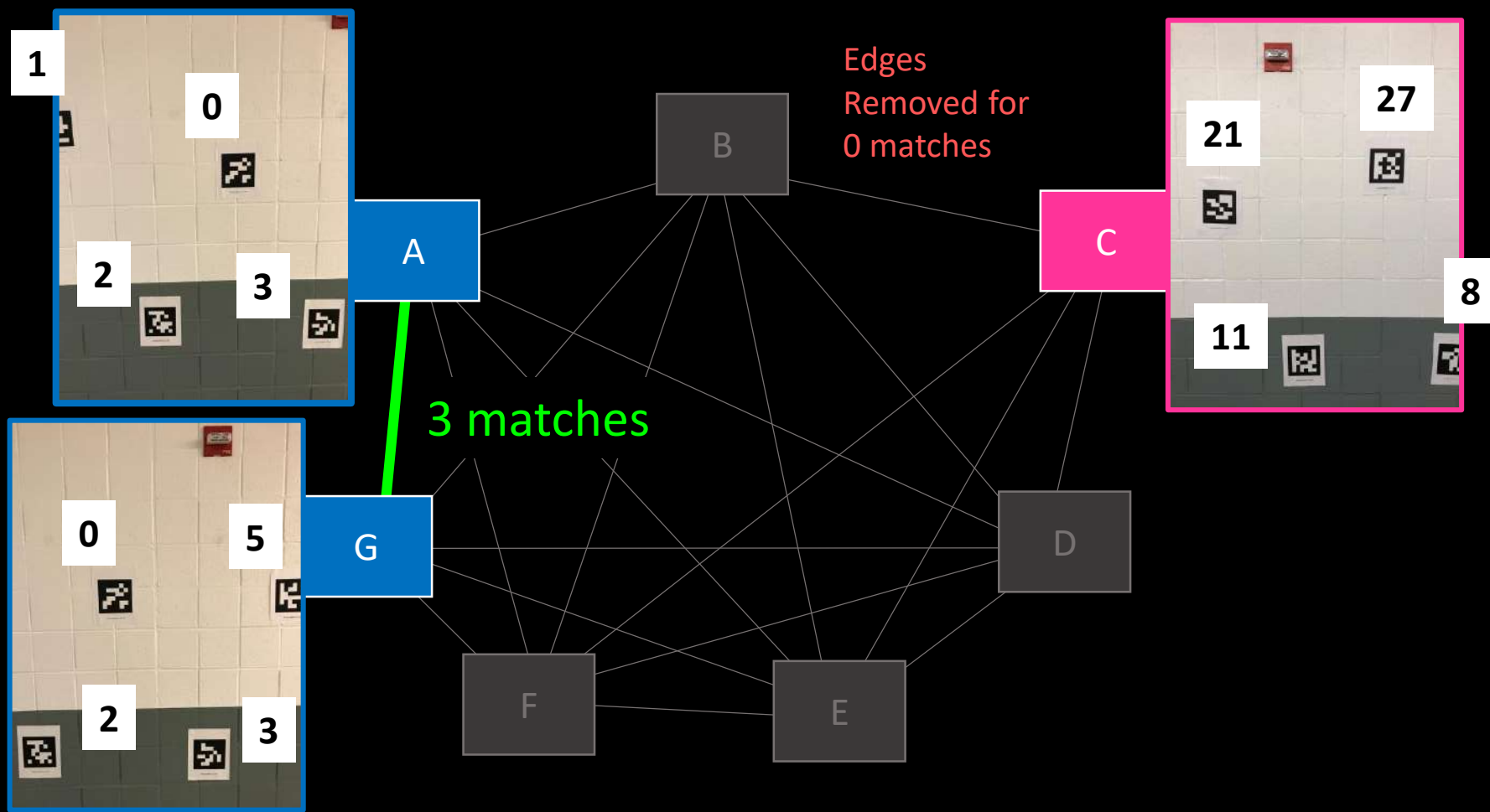
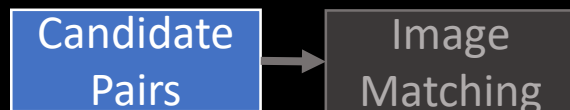
21 possible
image pairs



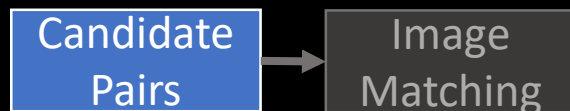
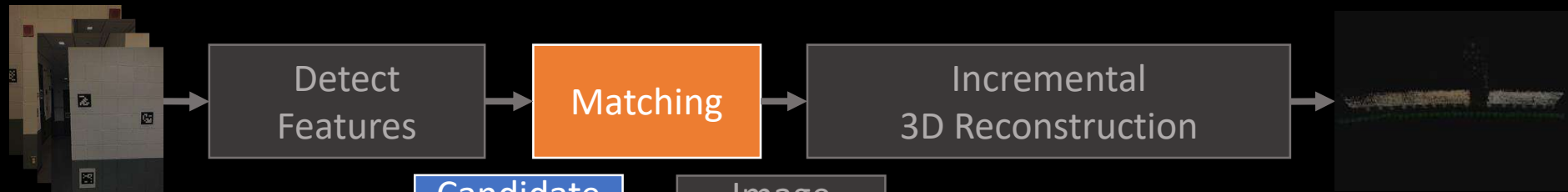
Structure from Motion Pipeline



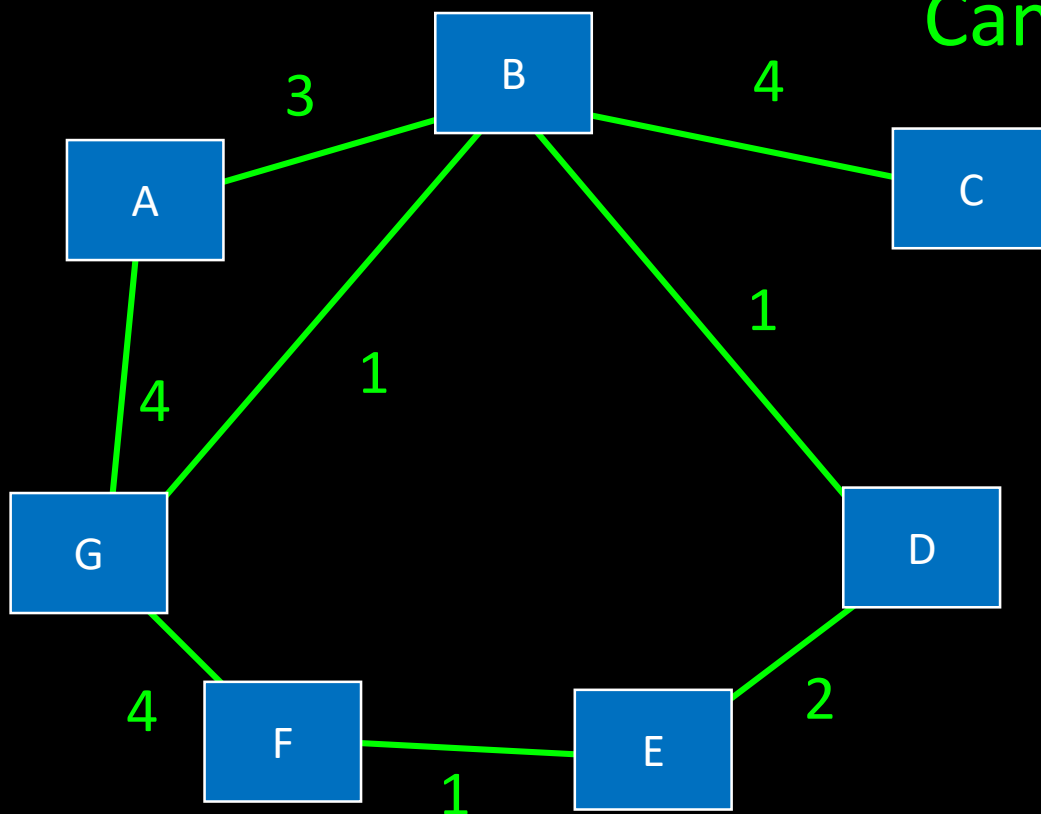
Structure from Motion Pipeline



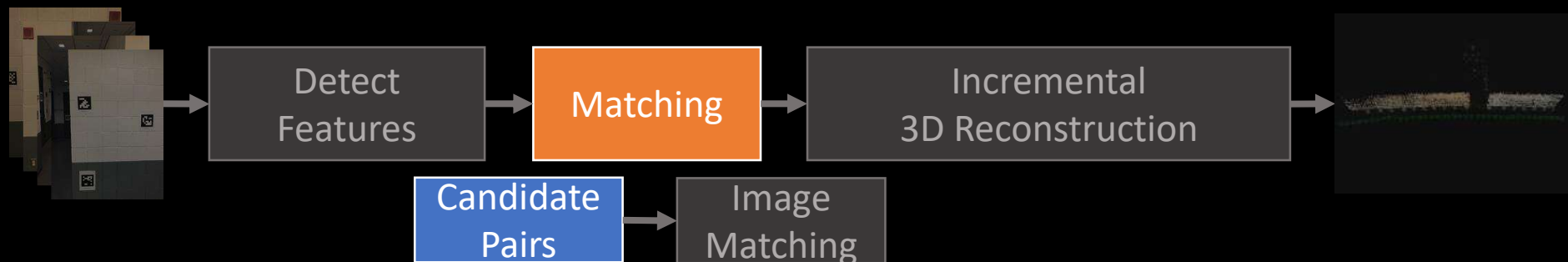
Structure from Motion Pipeline



8 Remaining
Candidate Pairs

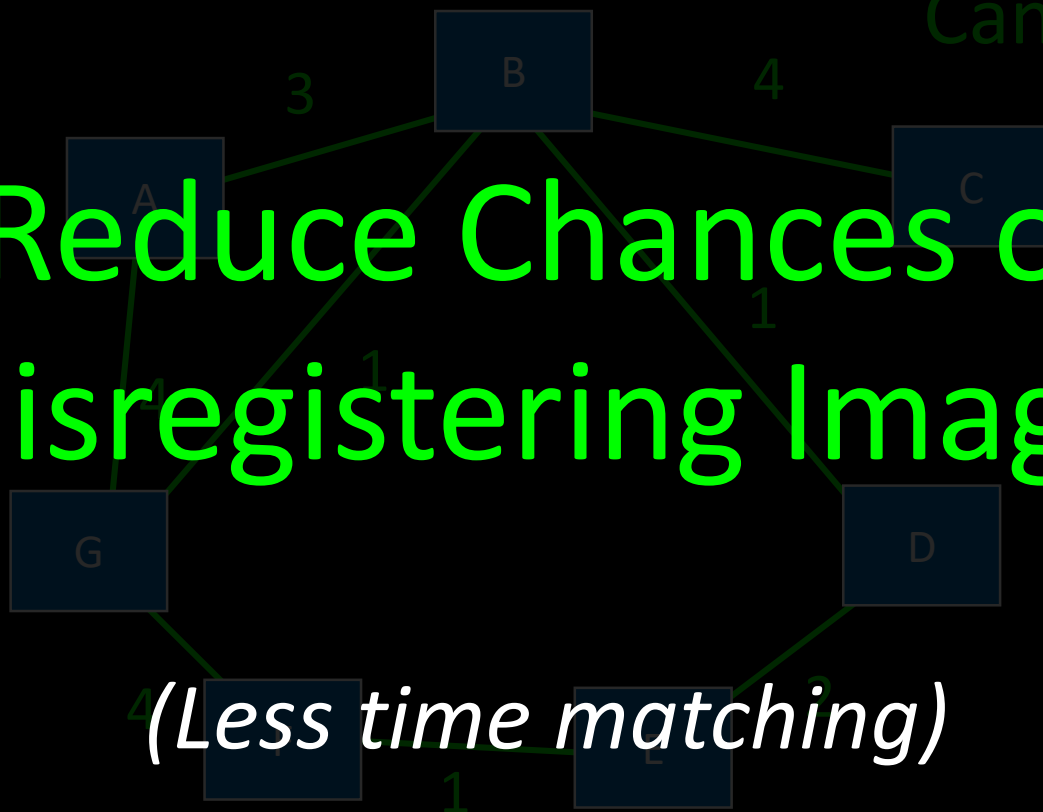


Structure from Motion Pipeline



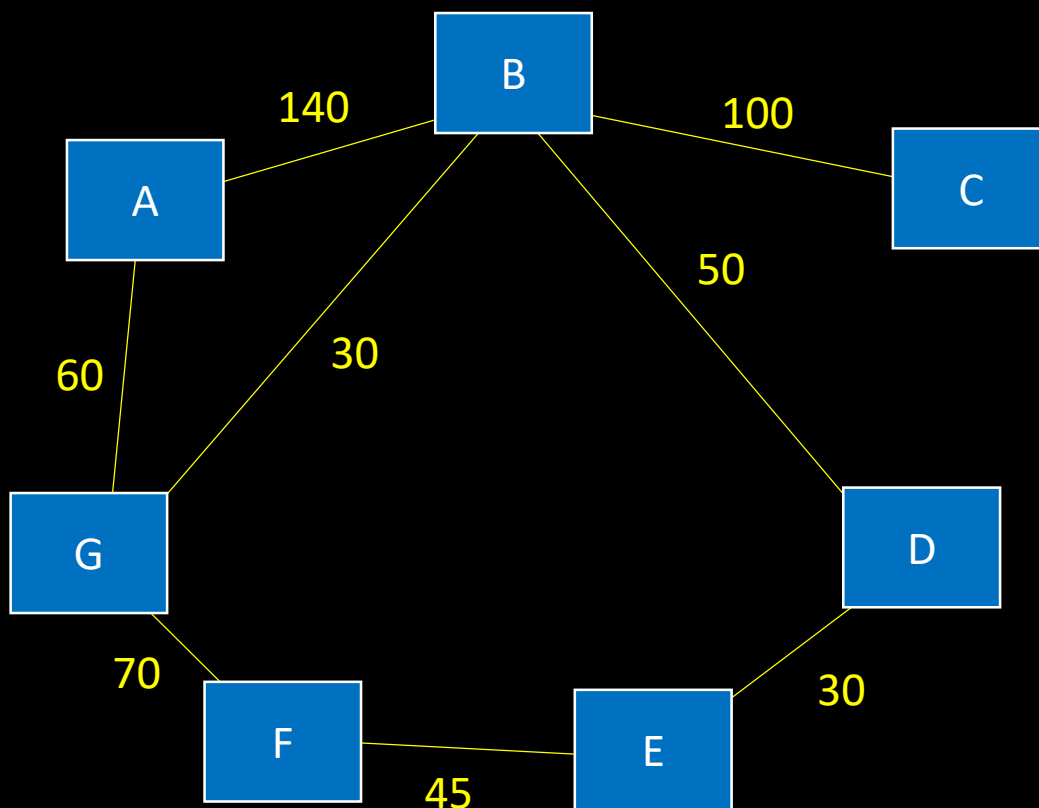
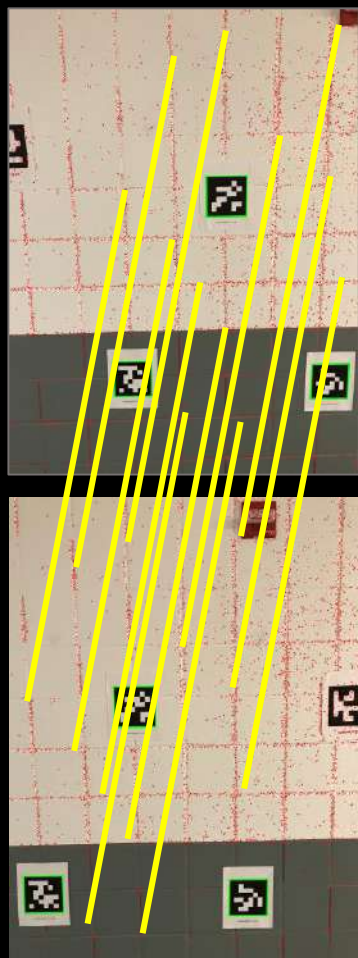
8 Remaining
Candidate Pairs

Reduce Chances of
Misregistering Images

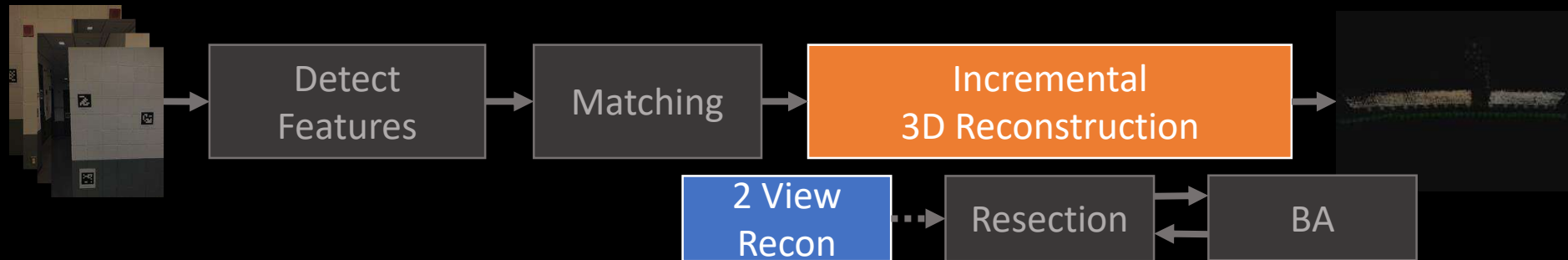


(Less time matching)

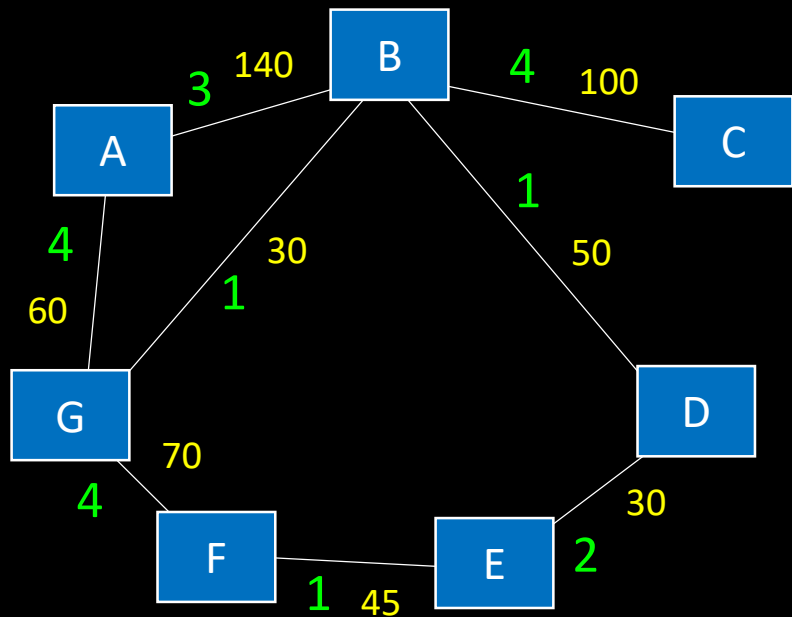
Structure from Motion Pipeline



Structure from Motion Pipeline



Match Graph



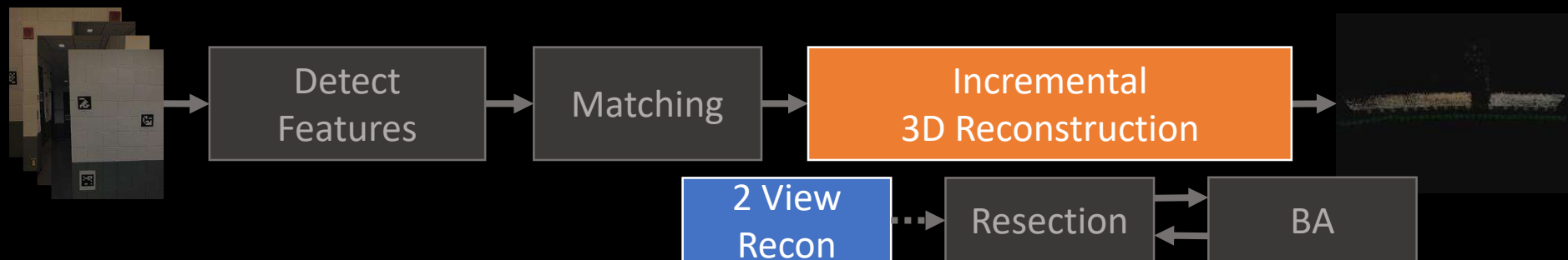
Marker Matches

Image Feature Matches

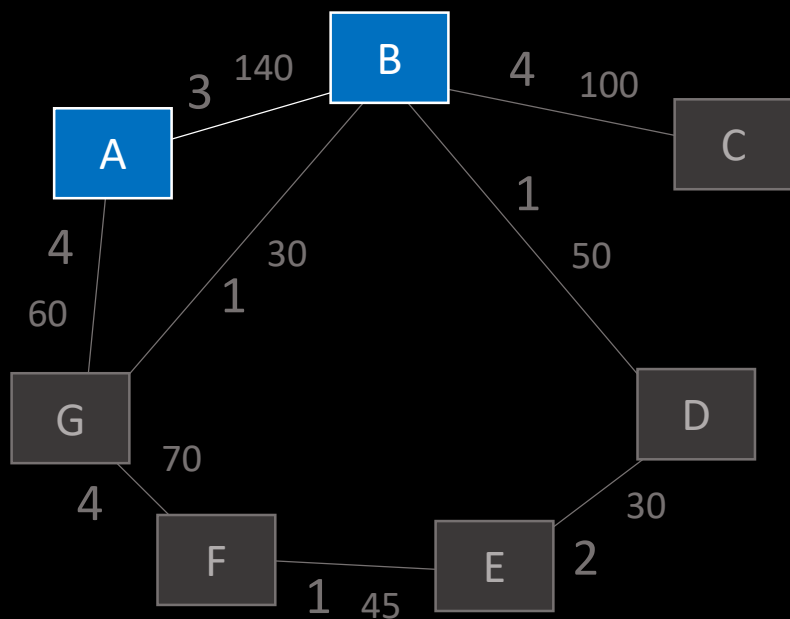
3D Reconstruction

Need to start the reconstruction...

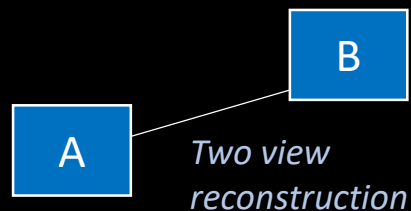
Structure from Motion Pipeline



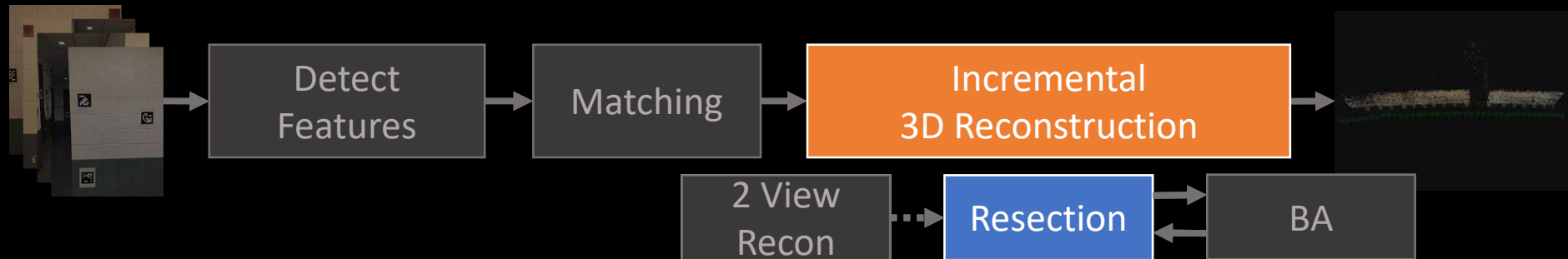
Match Graph



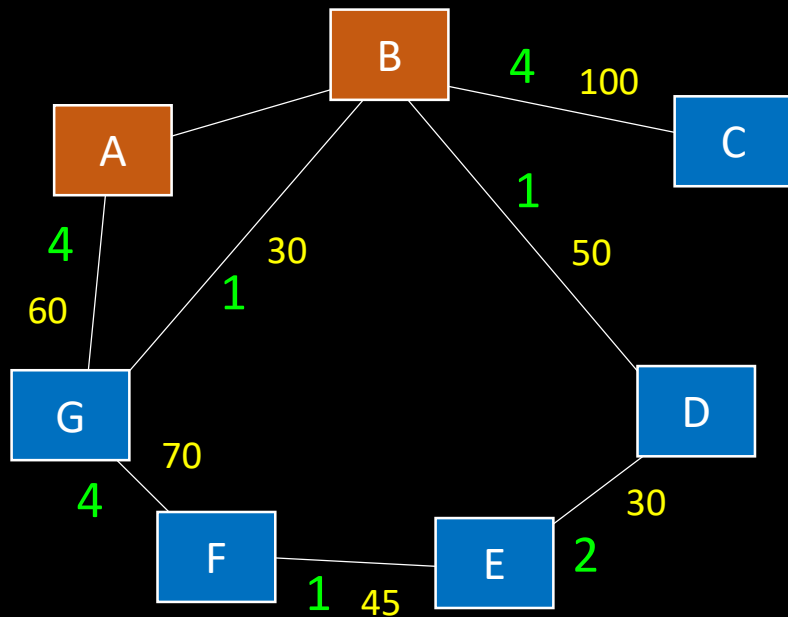
3D Reconstruction



Structure from Motion Pipeline



Match Graph

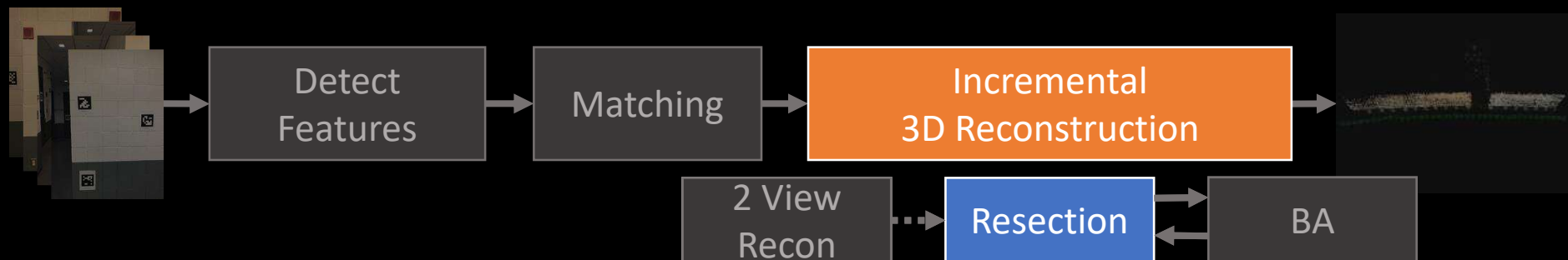


3D Reconstruction

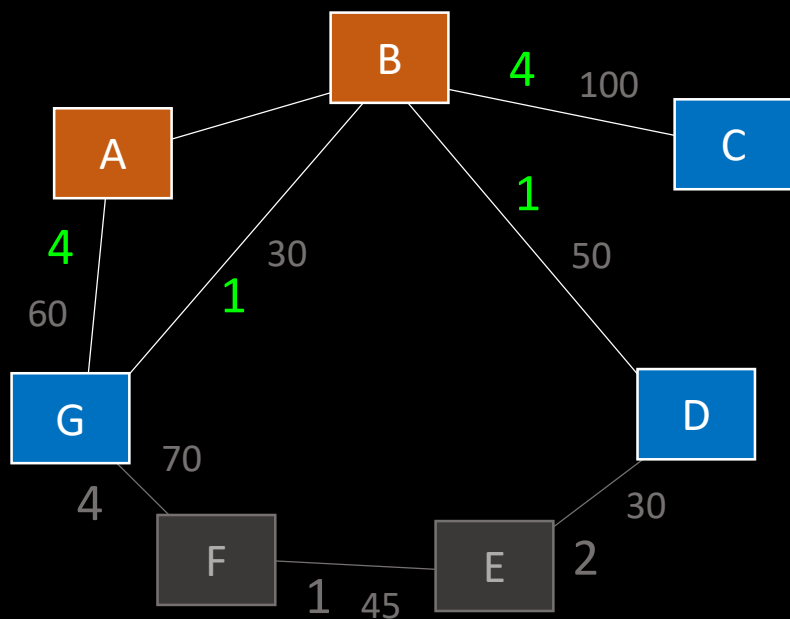


Need to add images to the reconstruction...

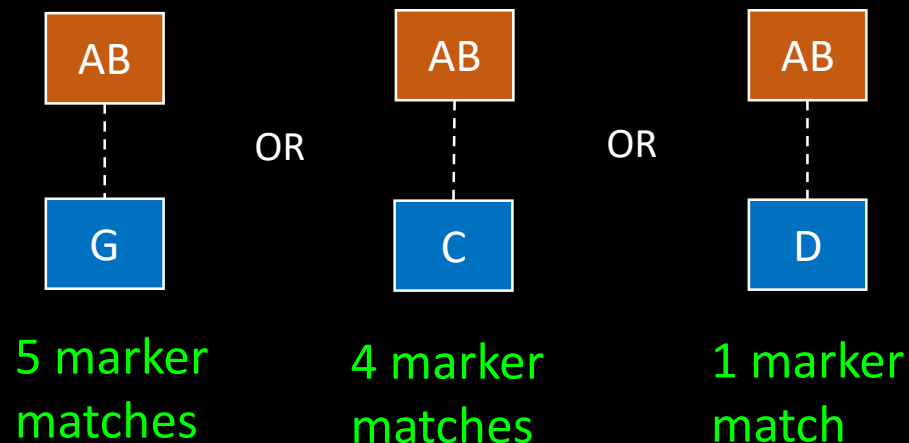
Structure from Motion Pipeline



Match Graph

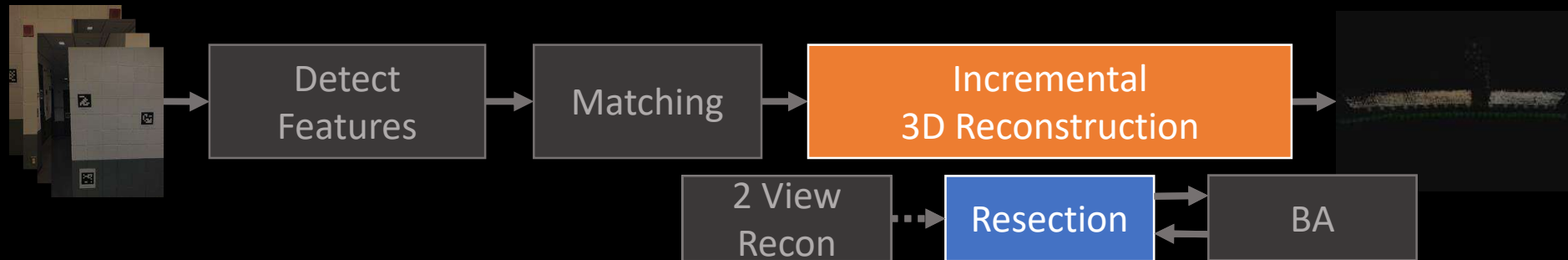


3D Reconstruction

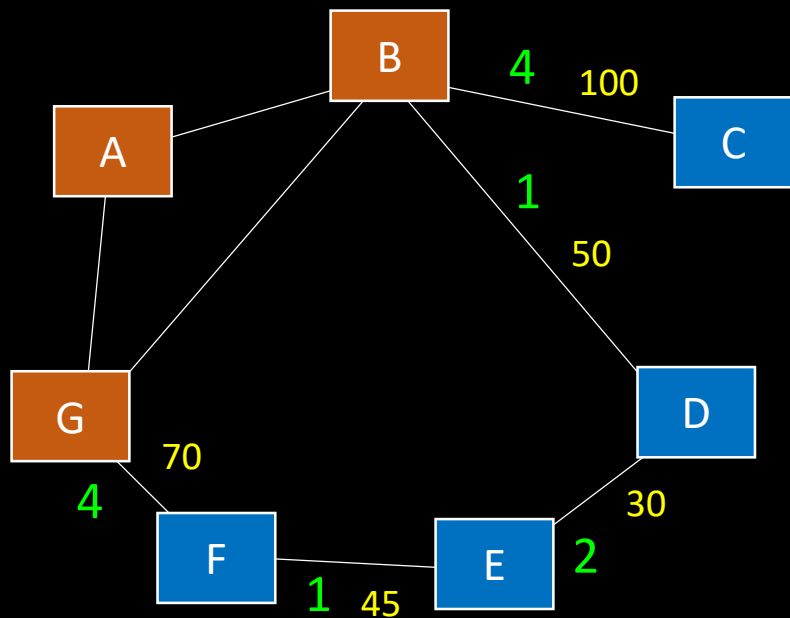


Rule 1: More Marker Matches Wins

Structure from Motion Pipeline



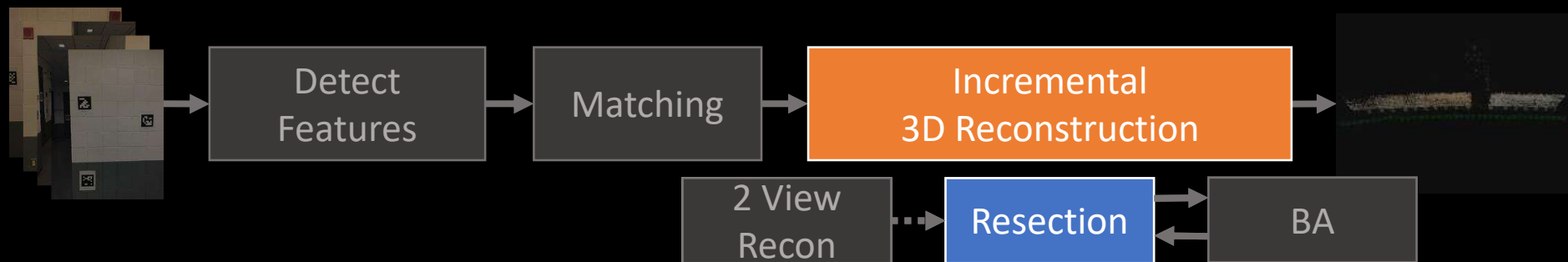
Match Graph



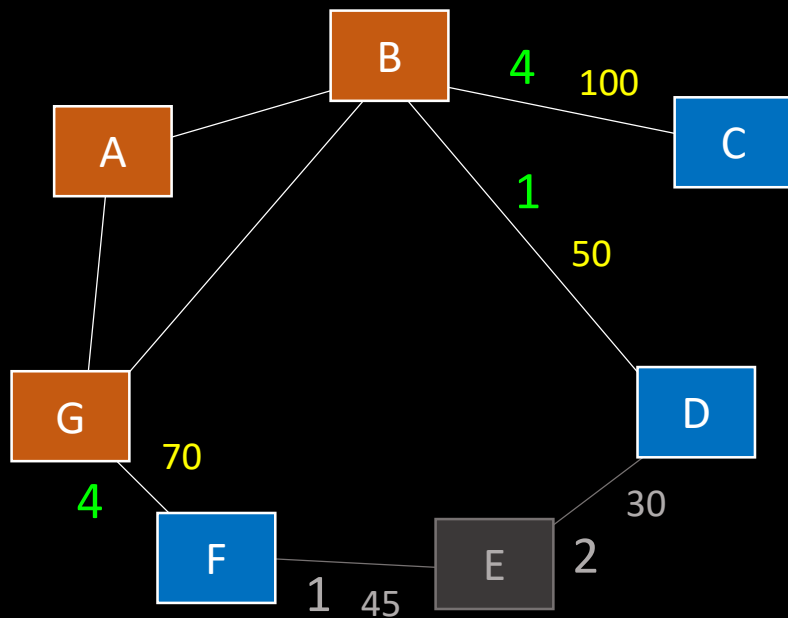
3D Reconstruction



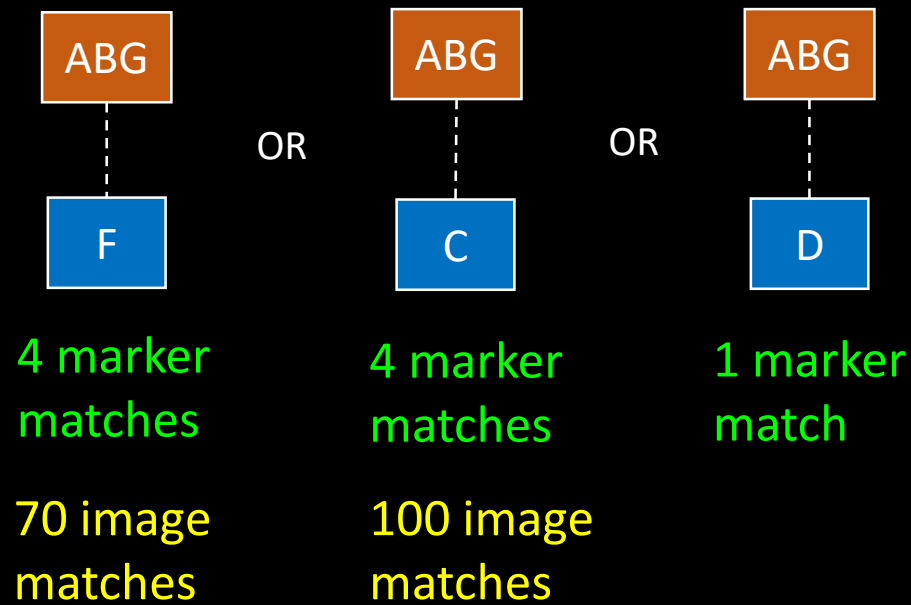
Structure from Motion Pipeline



Match Graph

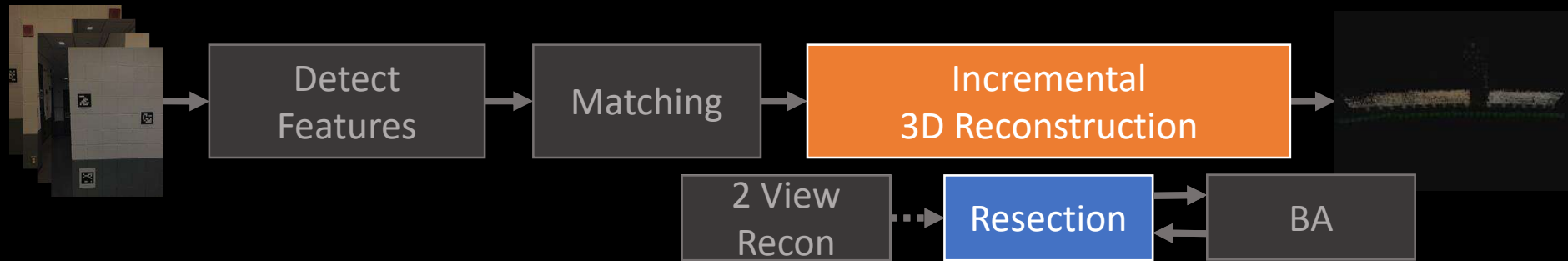


3D Reconstruction

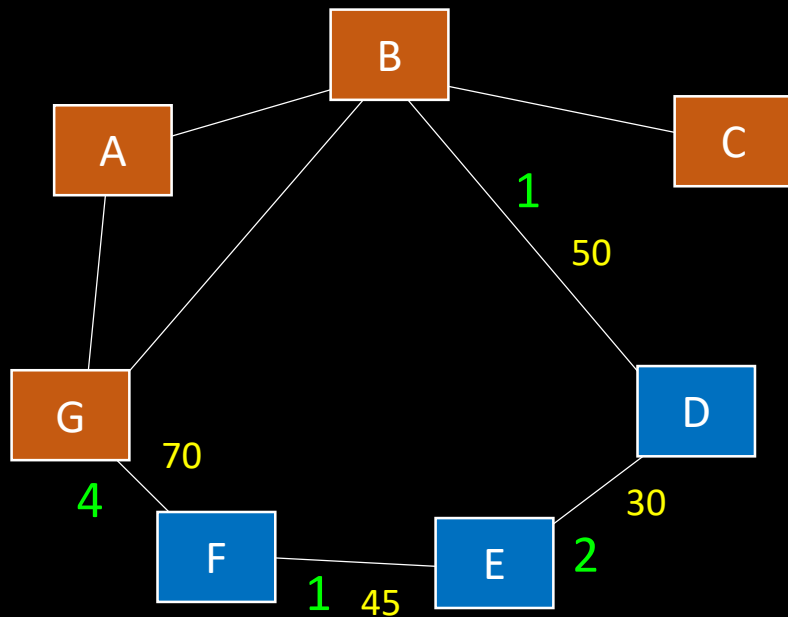


Rule 2: Feature Matches Break Ties

Structure from Motion Pipeline



Match Graph

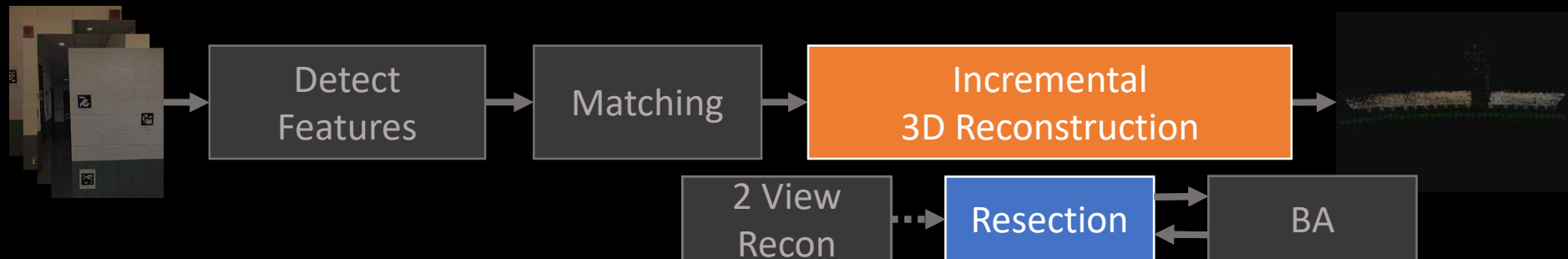


3D Reconstruction

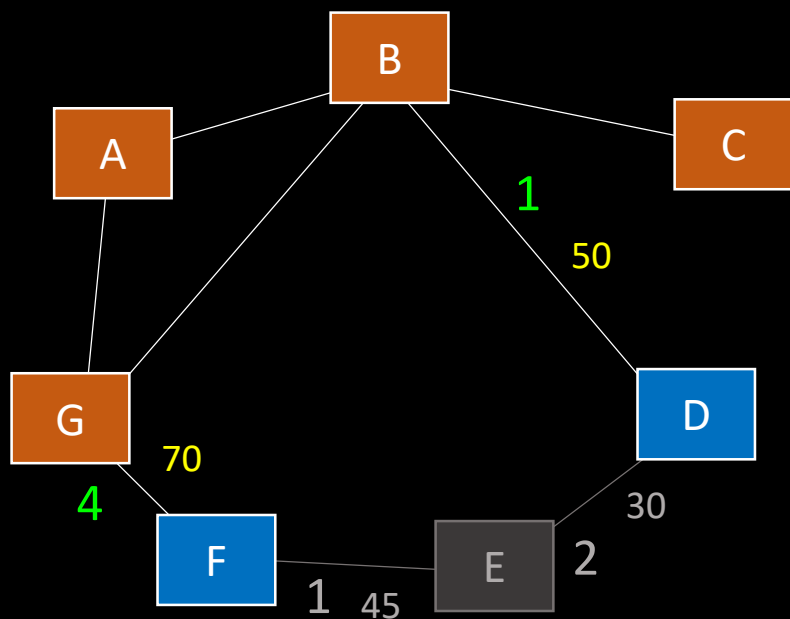


Continue this process to resection all images...

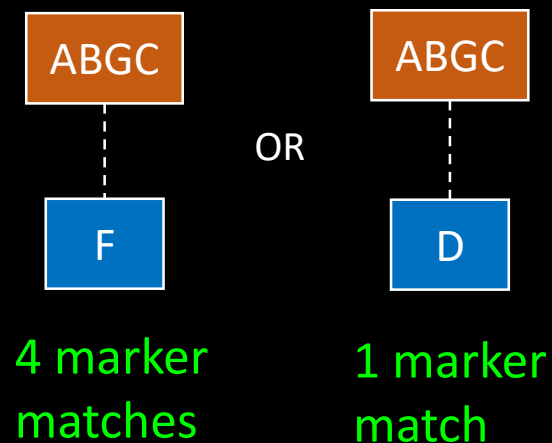
Structure from Motion Pipeline



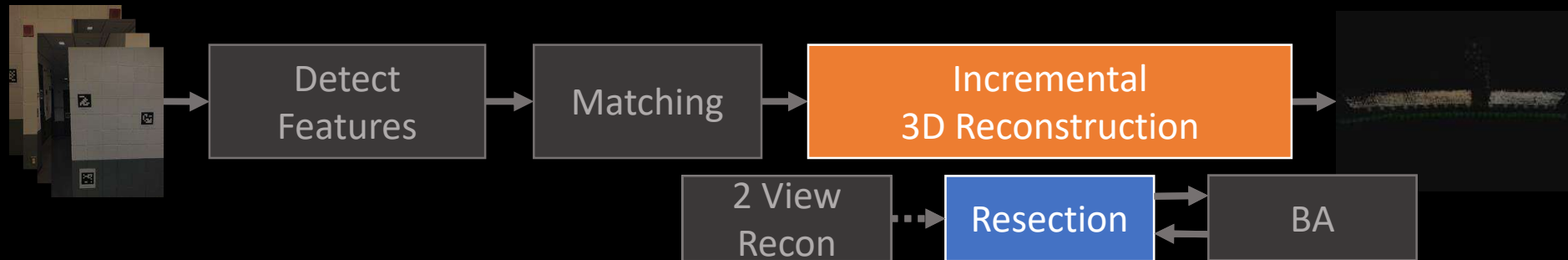
Match Graph



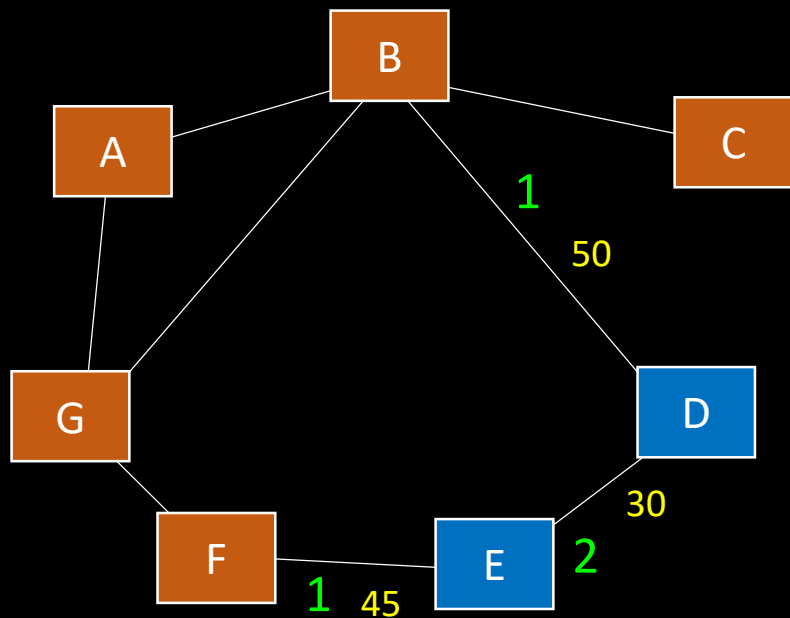
3D Reconstruction



Structure from Motion Pipeline



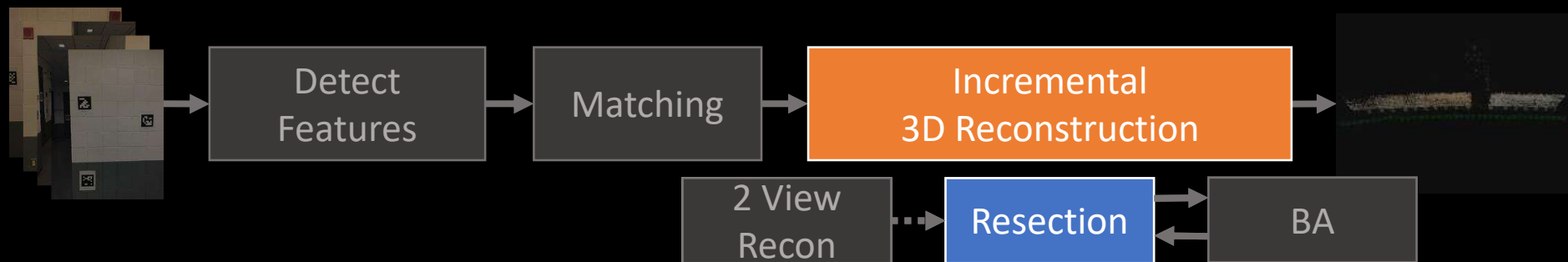
Match Graph



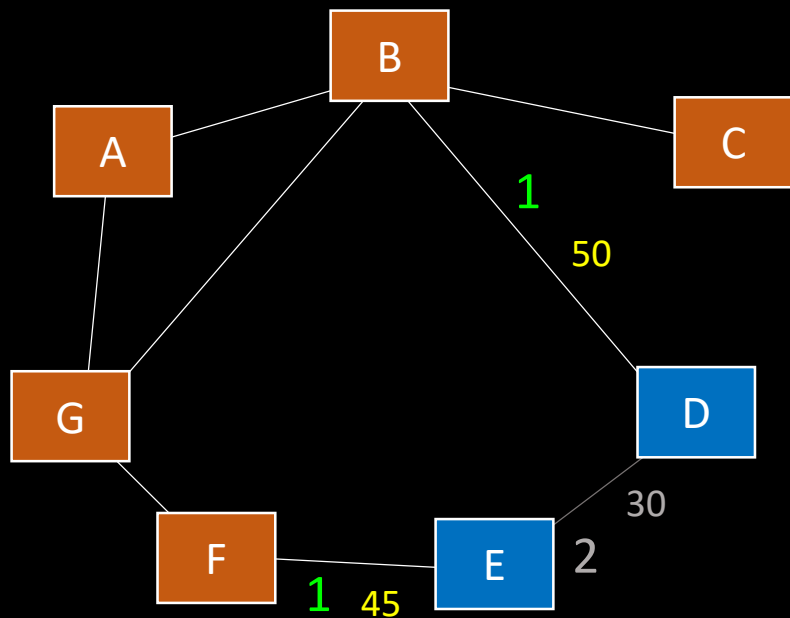
3D Reconstruction



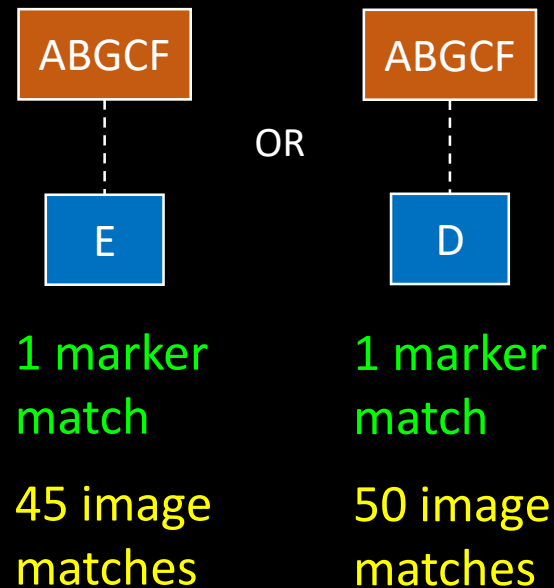
Structure from Motion Pipeline



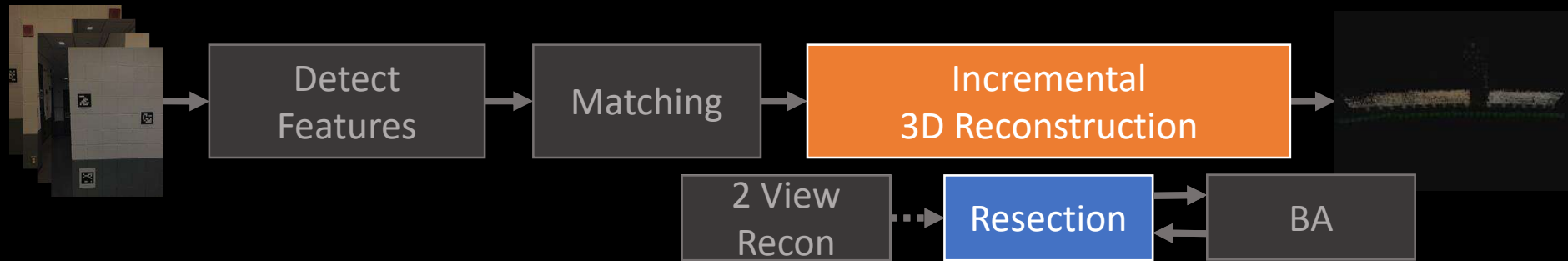
Match Graph



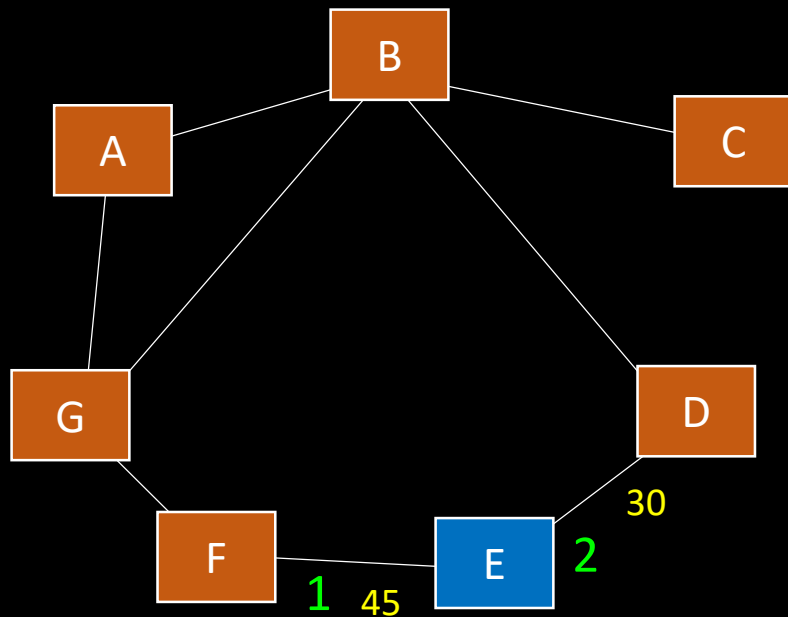
3D Reconstruction



Structure from Motion Pipeline



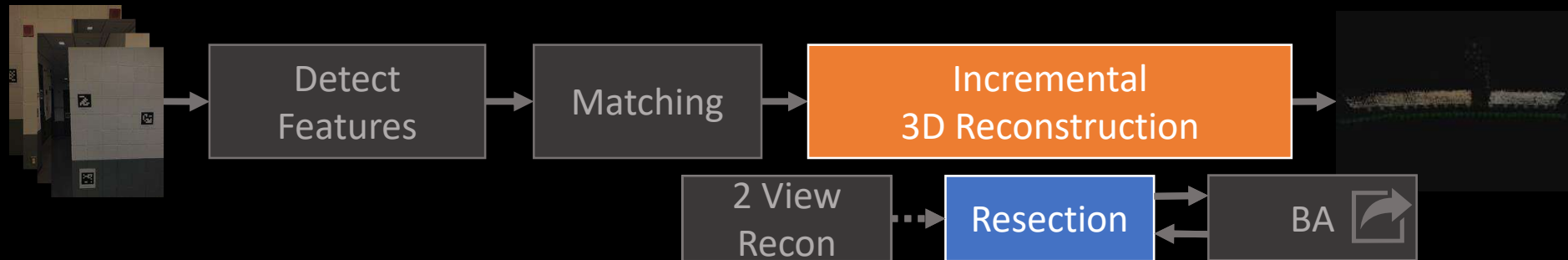
Match Graph



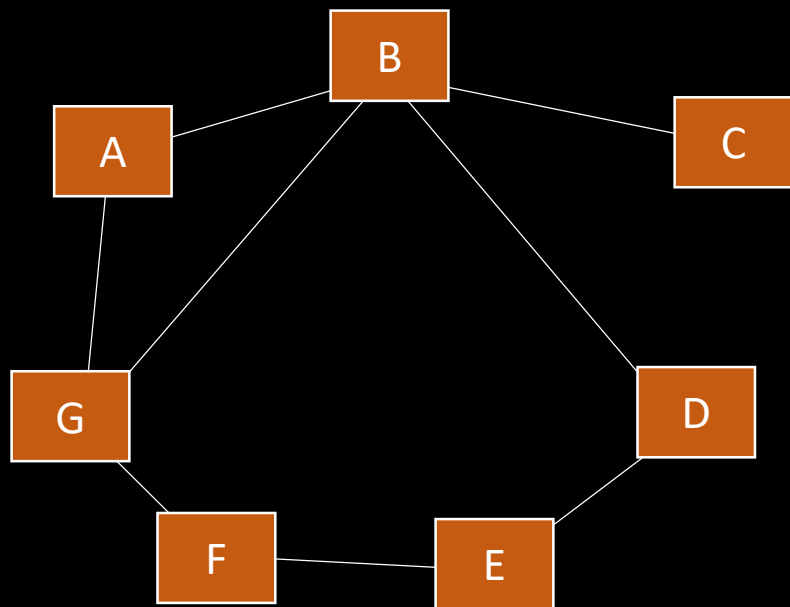
3D Reconstruction



Structure from Motion Pipeline



Match Graph



3D Reconstruction

ABGCFDE

Improved Structure from Motion Using Fiducial Marker Matching

Data and Results

Experimental Setup

Dataset

- 16 Indoor image collections (*3530 total images*)
- Tens of AprilTags used in smaller scenes
- Hundreds of AprilTags used in larger scenes
- 3 Buildings



ECEB



Newmark



MUF @ Research Park

Experimental Setup

Dataset

- 16 Indoor image collections (*3530 total images*)
- Tens of AprilTags used in smaller scenes
- Hundreds of AprilTags used in larger scenes
- 3 Buildings



ECEB



Newmark

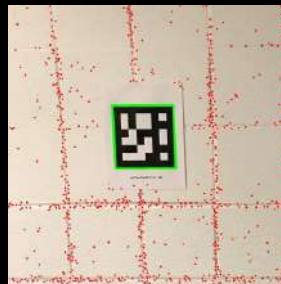


MUF @ Research Park

Methods for Comparison



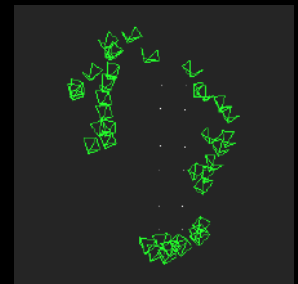
OpenSfM
by Mappilary



OpenSfM
+ Masking

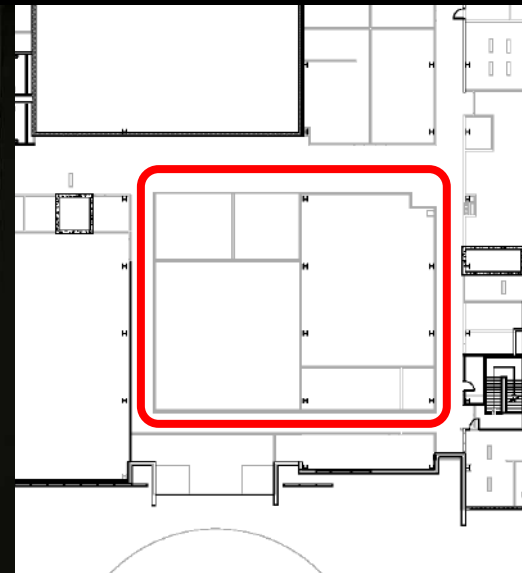
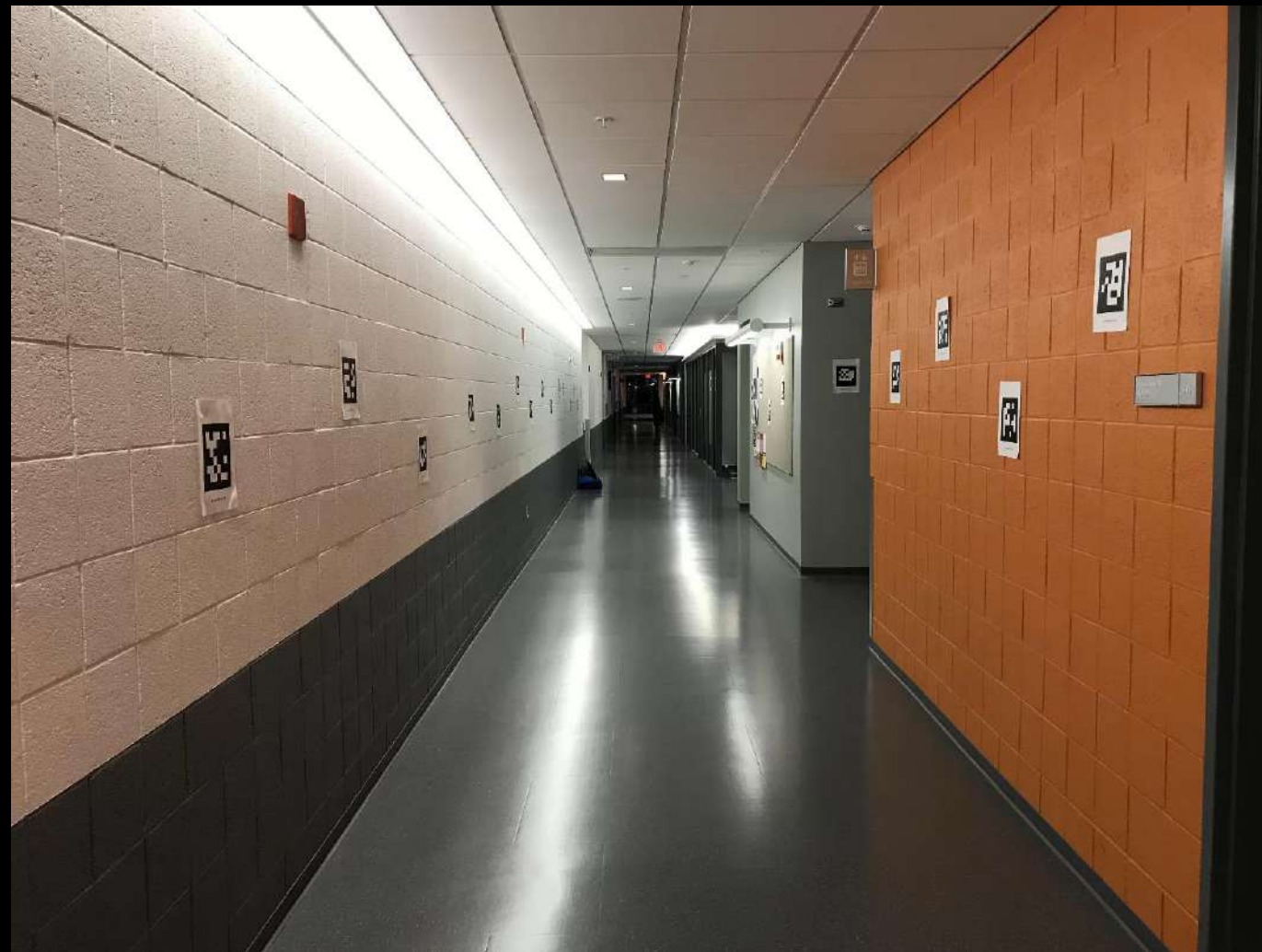


OpenSfM
+ Marker Tracks

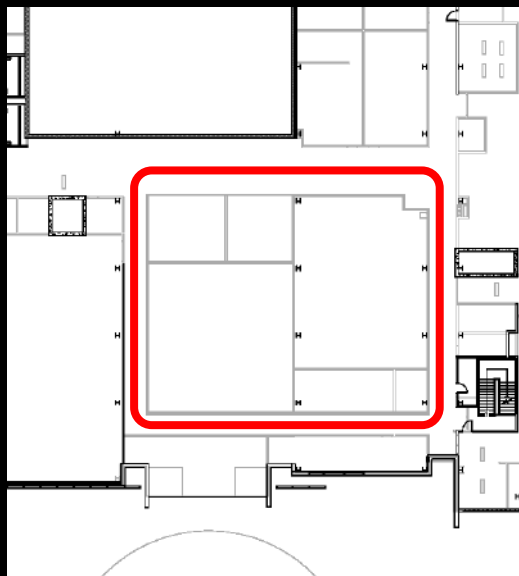


Marker Mapper
Pattern Recognition 2018

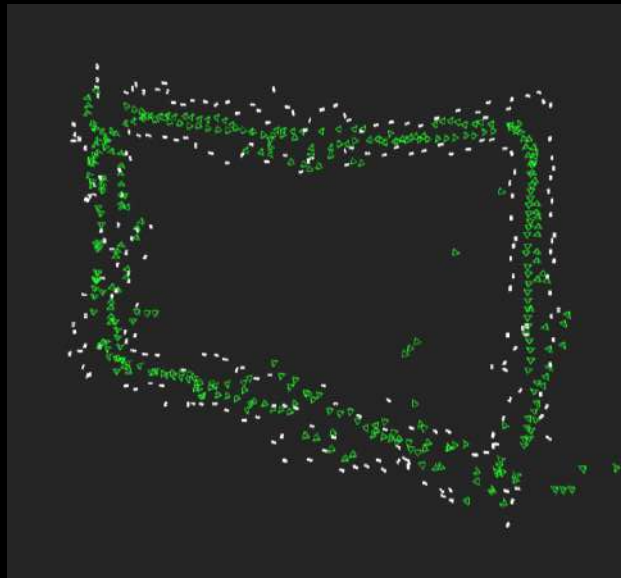
ECE Floor3 Loop



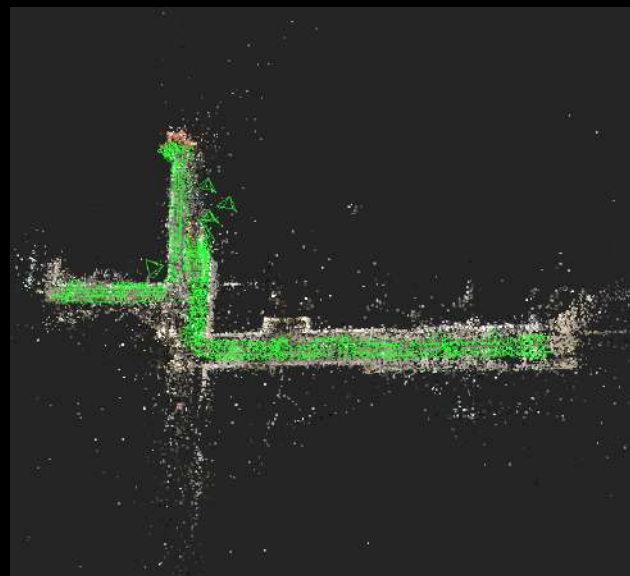
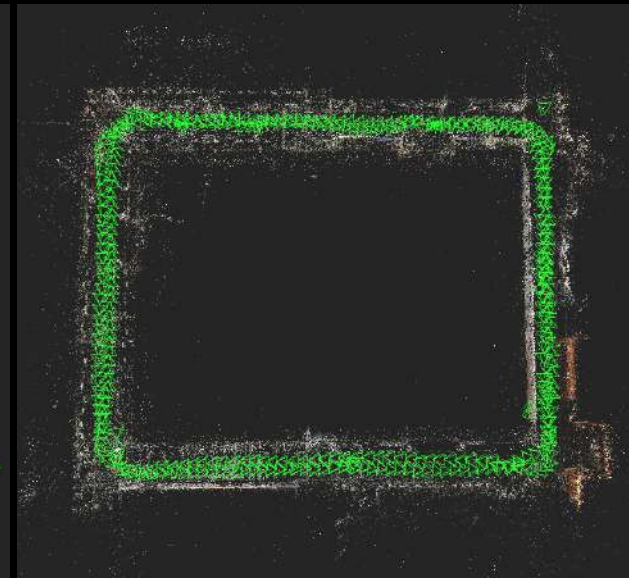
ECE Floor3 Loop



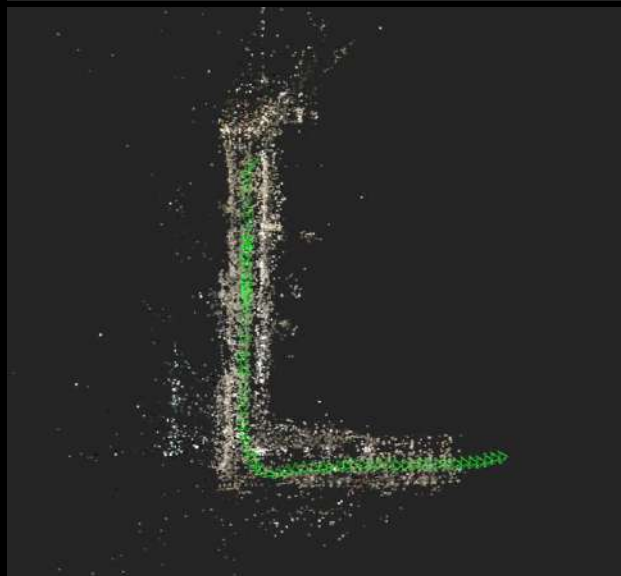
MarkerMapper



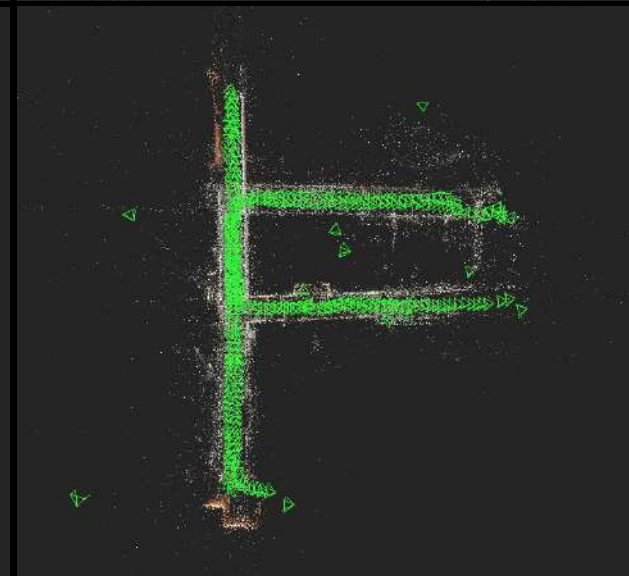
Our Method



OpenSfM

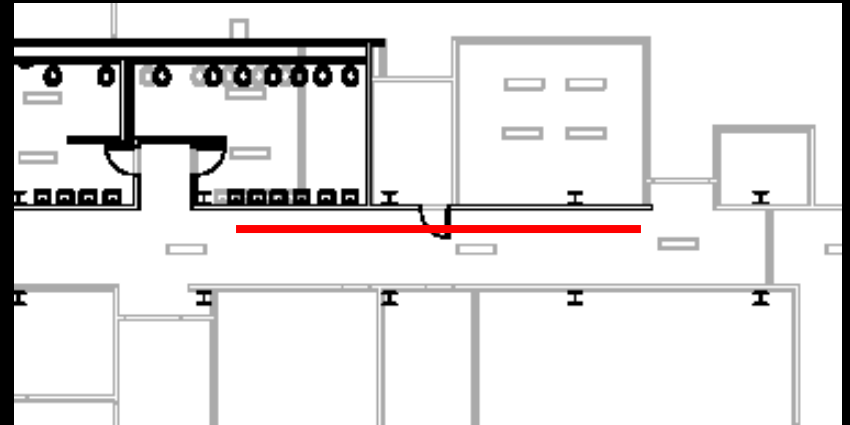


OpenSfM + Masking

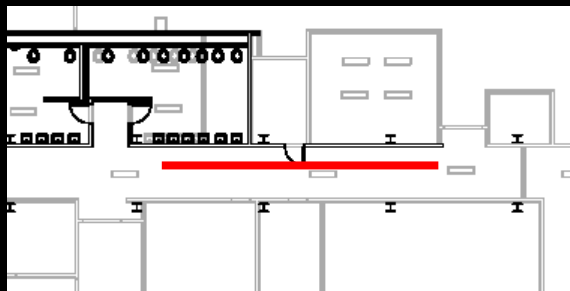


OpenSfM + Marker Tracks

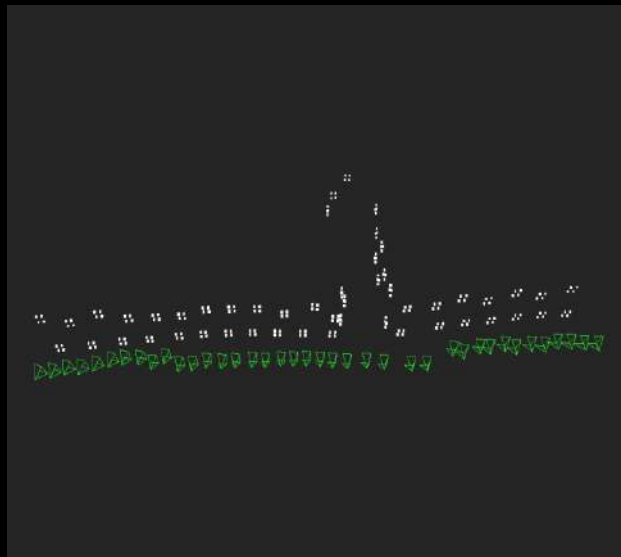
ECE Floor4 Wall



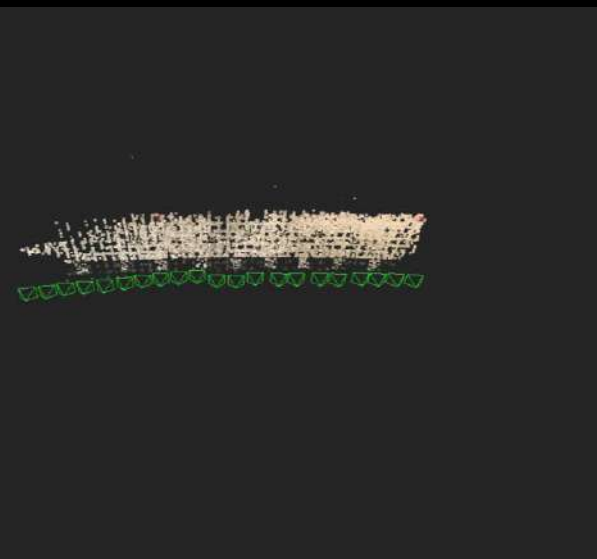
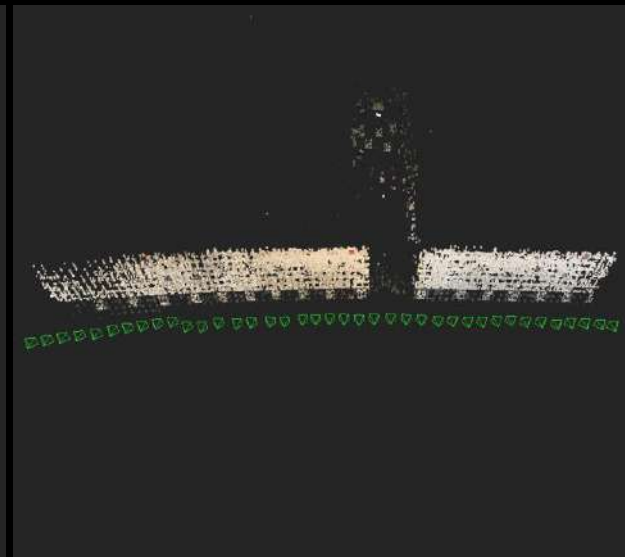
ECE Floor4 Wall



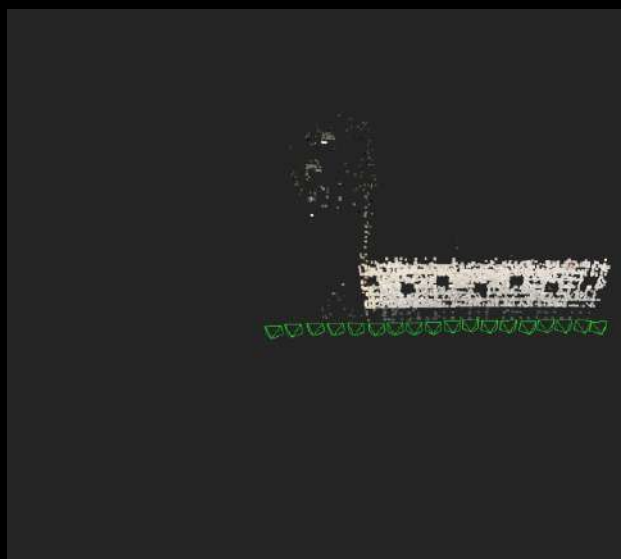
MarkerMapper



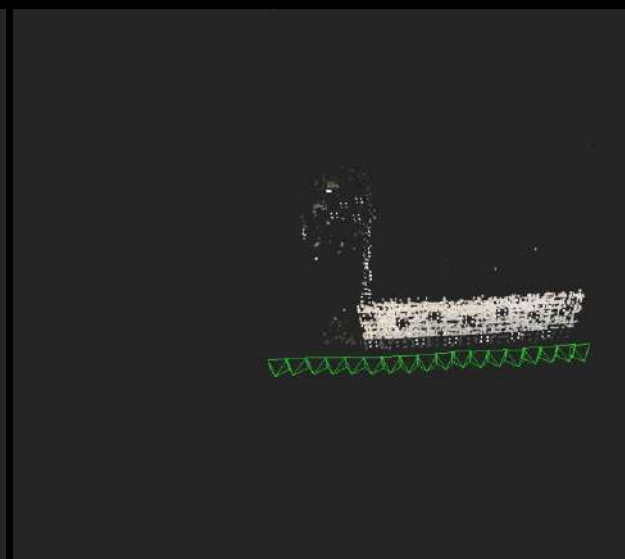
Our Method



OpenSfM

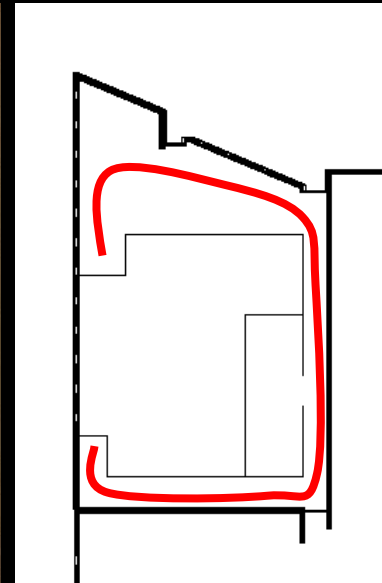


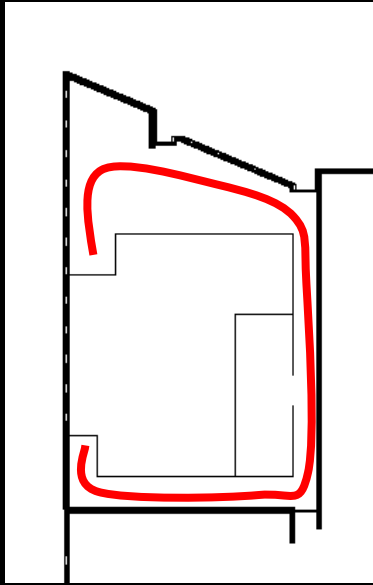
OpenSfM + Masking



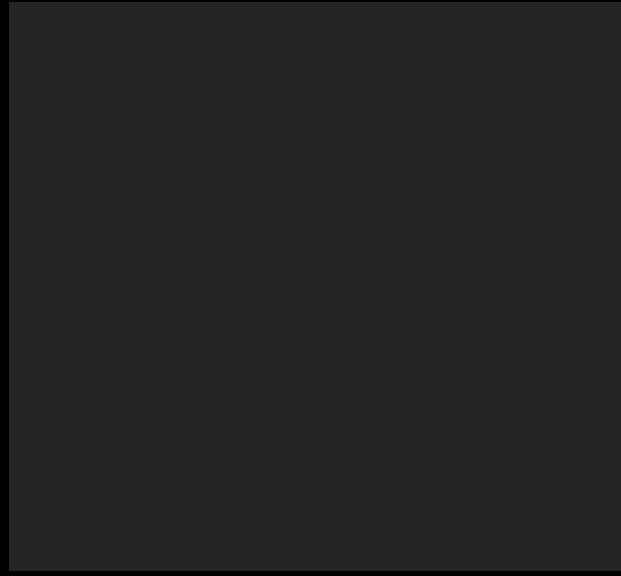
OpenSfM + Marker Tracks

CEE Day CCW

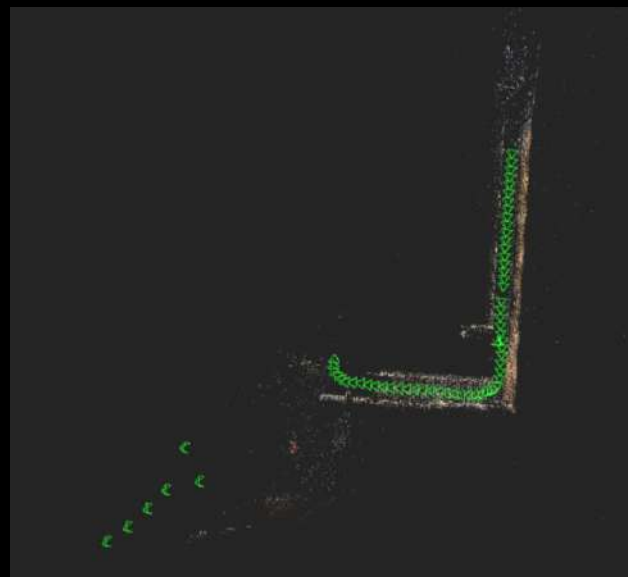
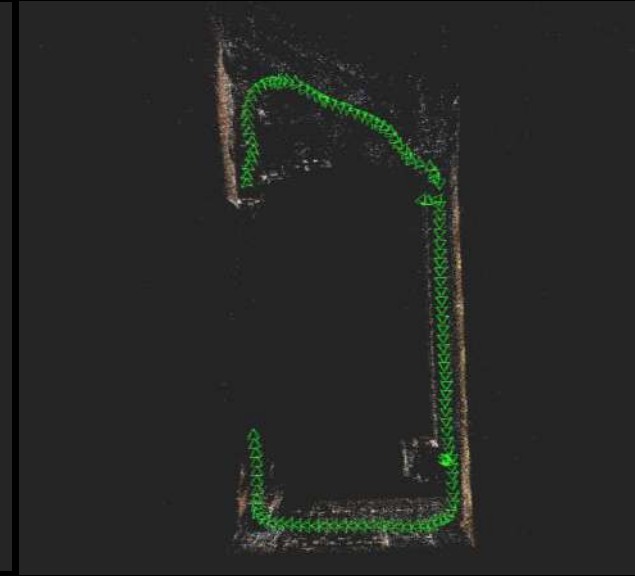




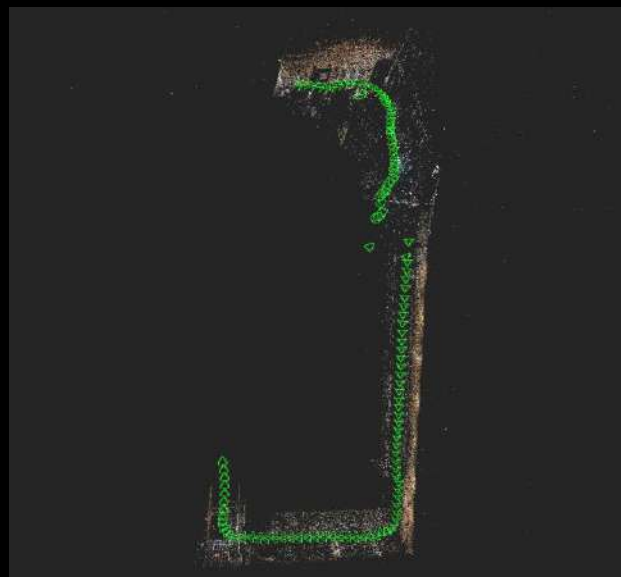
MarkerMapper



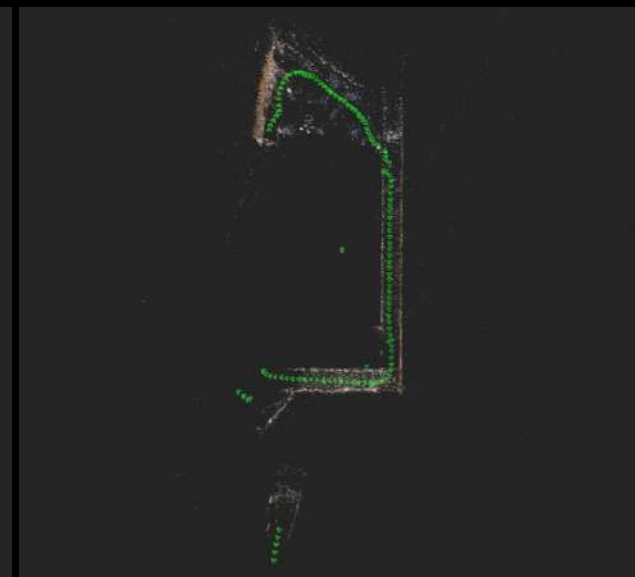
Our Method



OpenSfM

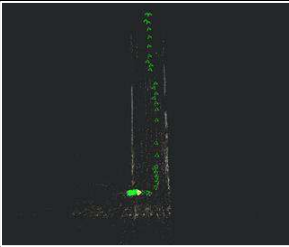
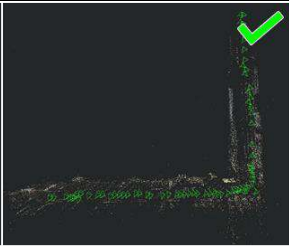
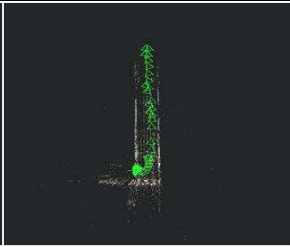
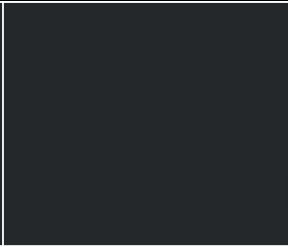
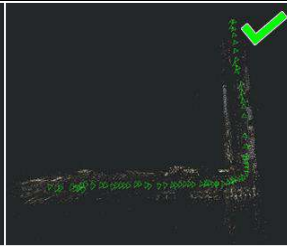
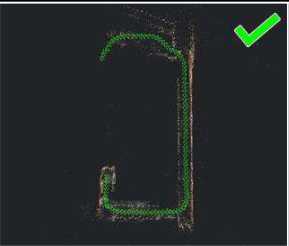
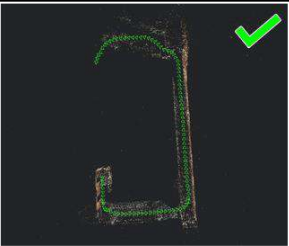

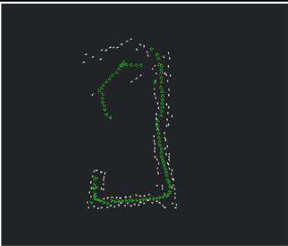
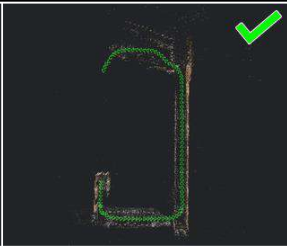







OpenSfM + Masking



OpenSfM + Marker Tracks

Results for all Datasets

Images		Methods				
		OpenSfM	OpenSfM + Masking	OpenSfM + Marker Tracks	MarkerMapper	Our Method
ECE F2 Hall	74					
ECE F3 Loop CCW	192					
ECE F3 Loop CW	170					
ECE F3 Loop	362					
ECE F5 Hall	239					
ECE Stairs	89					
ECE F5 Stairs	328					
ECE F4 Wall	39					
CEE Day CW	63					
CEE Day CCW	120					
CEE Day	252					
CEE Night CW	96					
CEE Night CCW	79					
CEE Night	170					
MUF F2	896					
MUF F3	361					

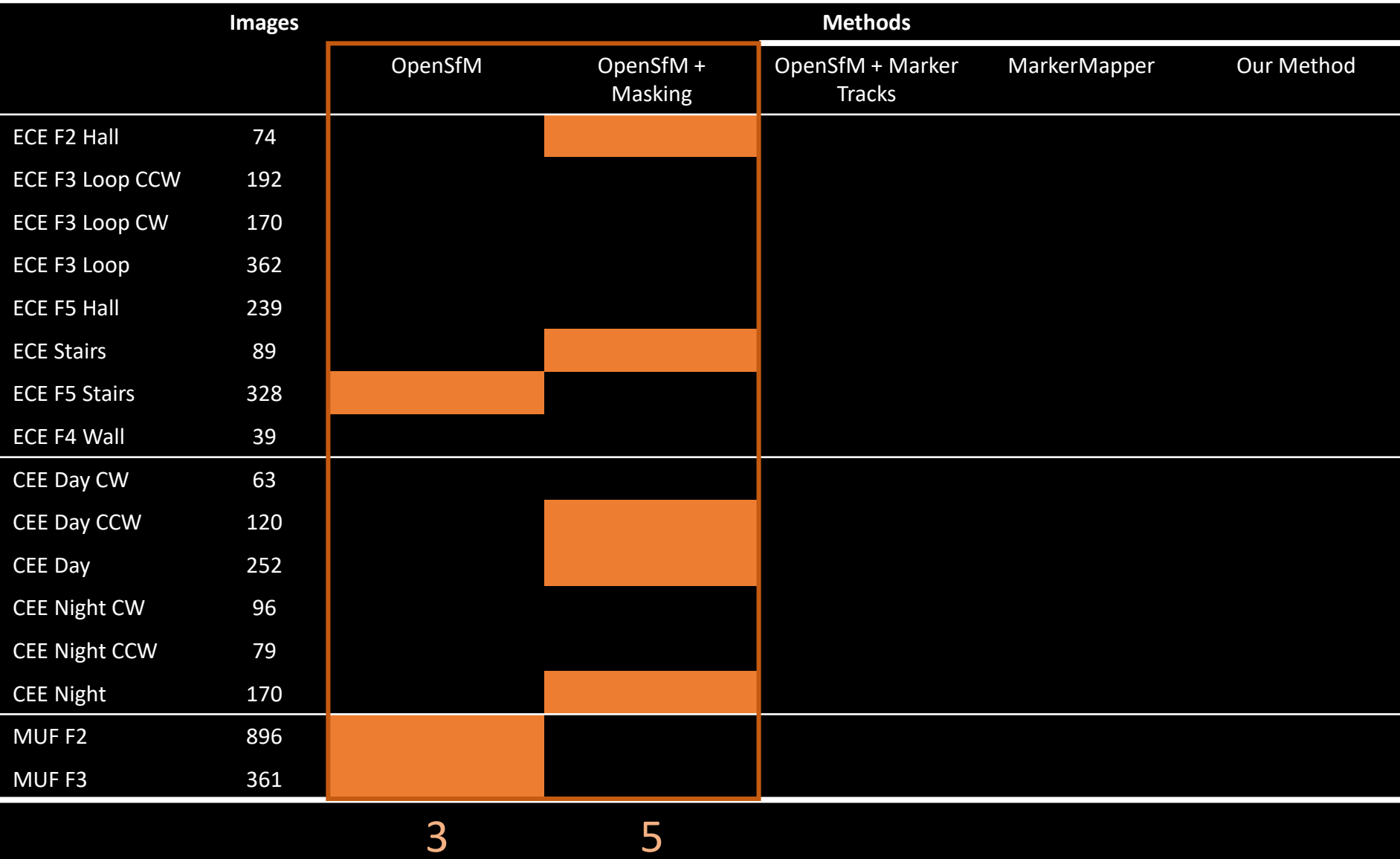
Results for all Datasets

Our method performs as well or better than all other methods on all datasets

Images		Methods				
		OpenSfM	OpenSfM + Masking	OpenSfM + Marker Tracks	MarkerMapper	Our Method
ECE F2 Hall	74					
ECE F3 Loop CCW	192					
ECE F3 Loop CW	170					
ECE F3 Loop	362					
ECE F5 Hall	239					
ECE Stairs	89					
ECE F5 Stairs	328					
ECE F4 Wall	39					
CEE Day CW	63					
CEE Day CCW	120					
CEE Day	252					
CEE Night CW	96					
CEE Night CCW	79					
CEE Night	170					
MUF F2	896					
MUF F3	361					
		3	4	2	4	16

Results for all Datasets

Masking markers is sometimes better than using markers as texture alone



Both Matching and Resectioning Matter

Both Matching and Resectioning improve results individually

	Mean % Images Registered
OpenSfM + Markers Masked	42.3 %
No Marker Informed Matching	49.8 %
No Marker Informed Resectioning	67.7 %
Full Method	98.4 %

Best results when both are used together

Vision Based Robots Monitoring Built Environments

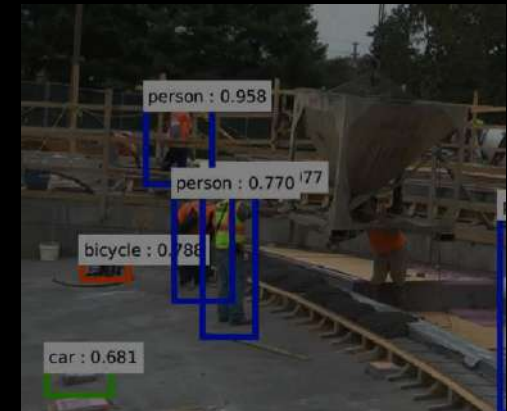
Data Capture



Mapping



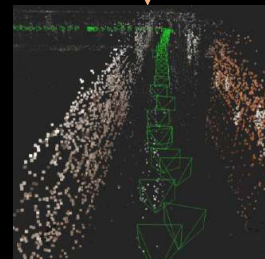
Analysis



Structure from Motion



Simulating Features to
Predict Mapping Failure



FEATS: Synthetic Feature Tracks for Structure from Motion Evaluation

*Joseph DeGol,
Jae Yong Lee,
Rajbir Kataria,
Daniel Yuan,
Timothy Bretl,
Derek Hoiem*

Submitted to ECCV 2018

Introduction

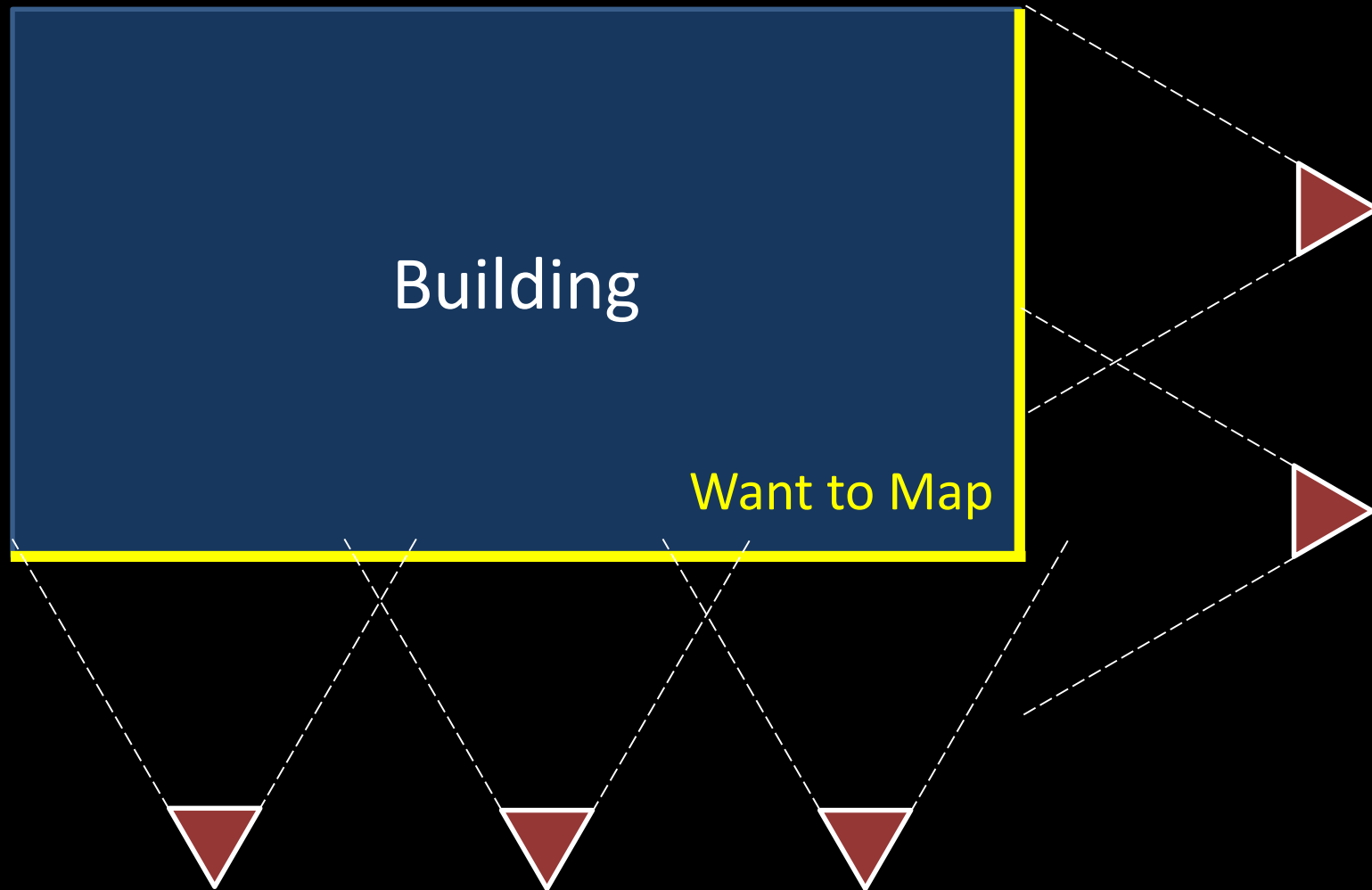
Feature
Extraction
And
Tracking
Simulator

Image Collection Paths Can Lead to Failure

Building

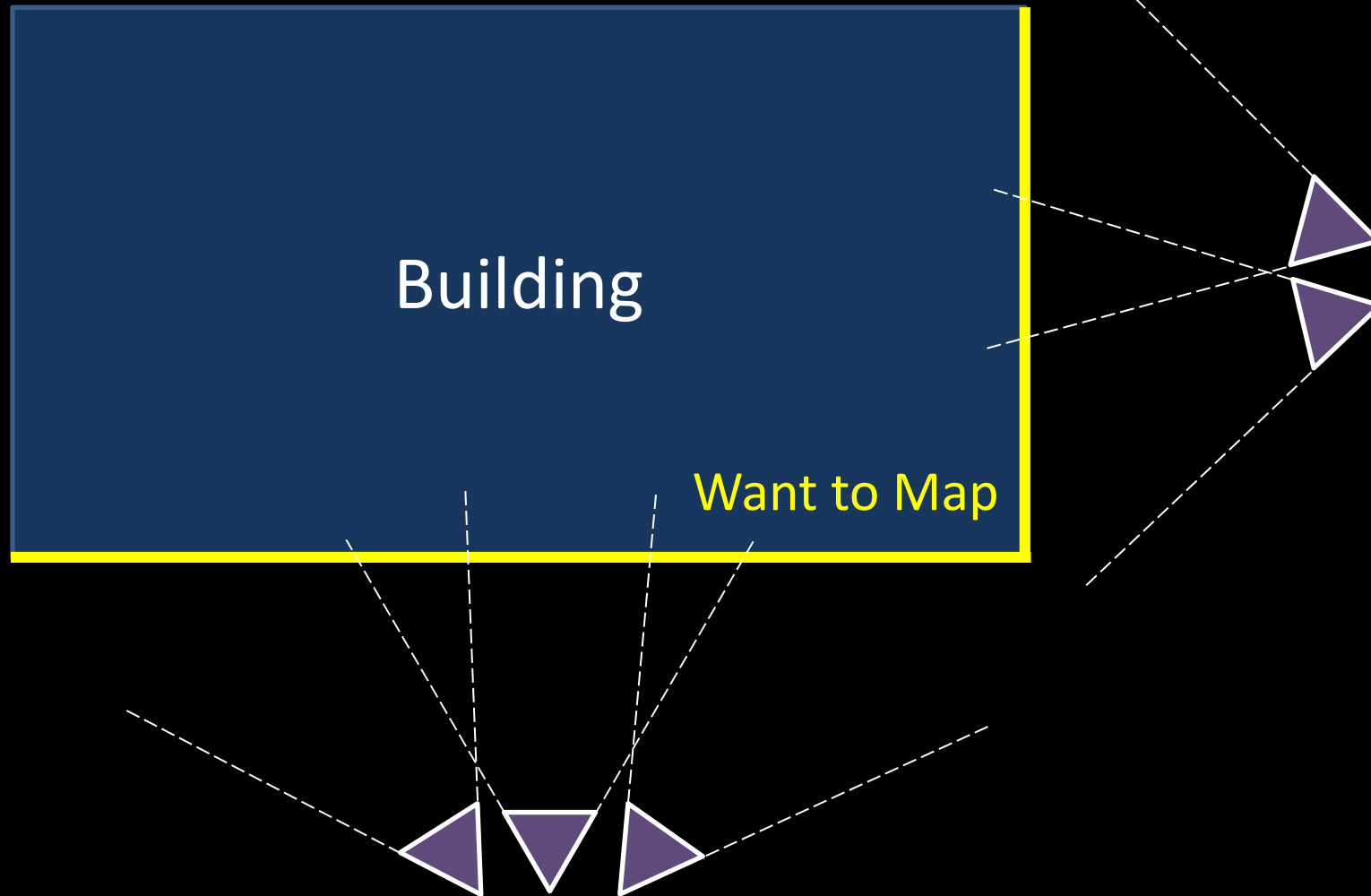
Want to Map

Image Collection Paths Can Lead to Failure



Not enough overlap between images

Image Collection Paths Can Lead to Failure

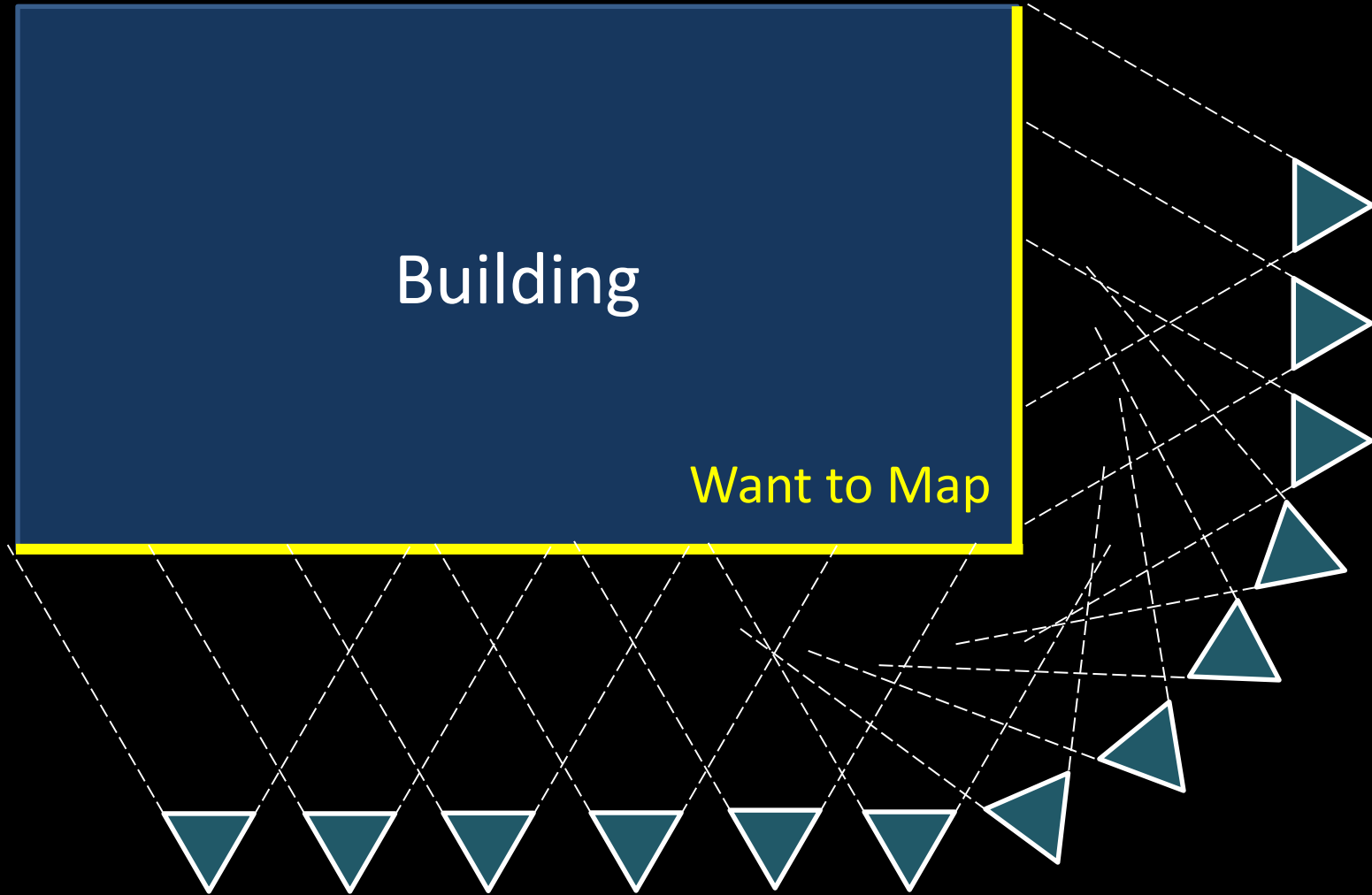


Not enough translation between images

Image Collection Paths Can Lead to Failure

Building

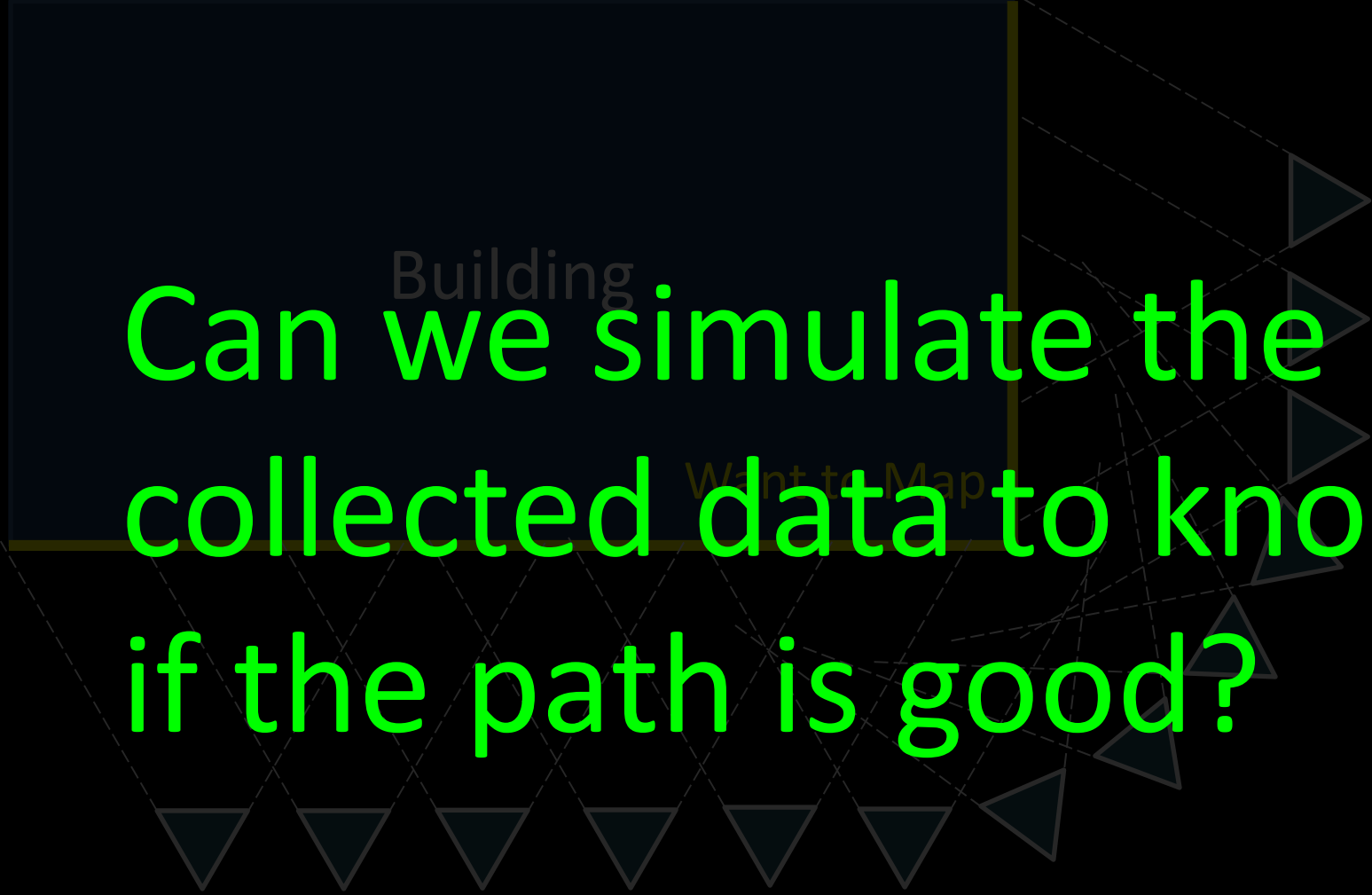
Want to Map



Intuition says this path is better

Do not know until after collection and processing

Image Collection Paths Can Lead to Failure



Building

Want to Map

Can we simulate the collected data to know if the path is good?

Intuition says this path is better

Do not know until after collection and processing

Simulating the Data Capture

Takes a user defined image path in a 3D map



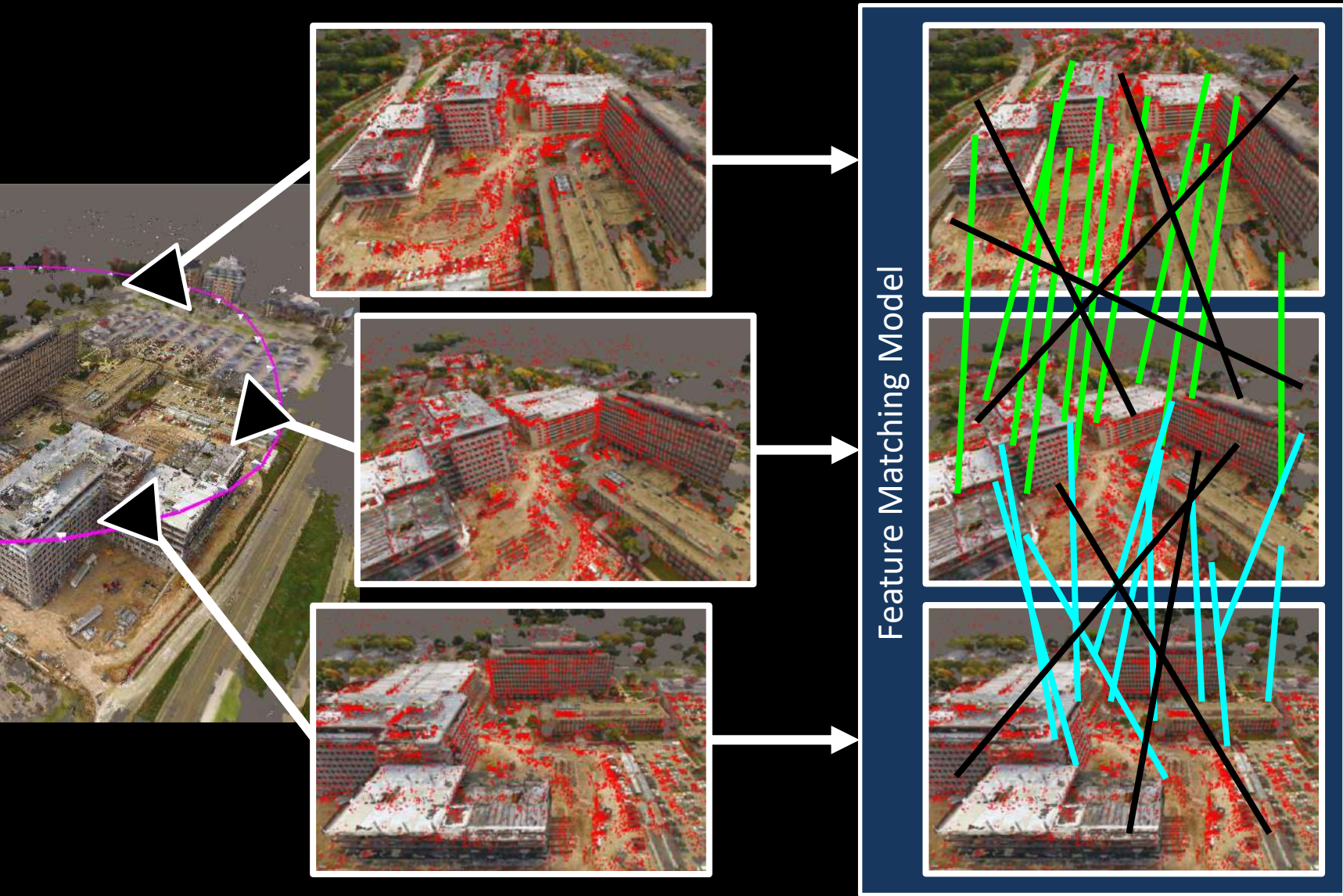
Simulating the Data Capture

Simulates Image Features for Images Along Path



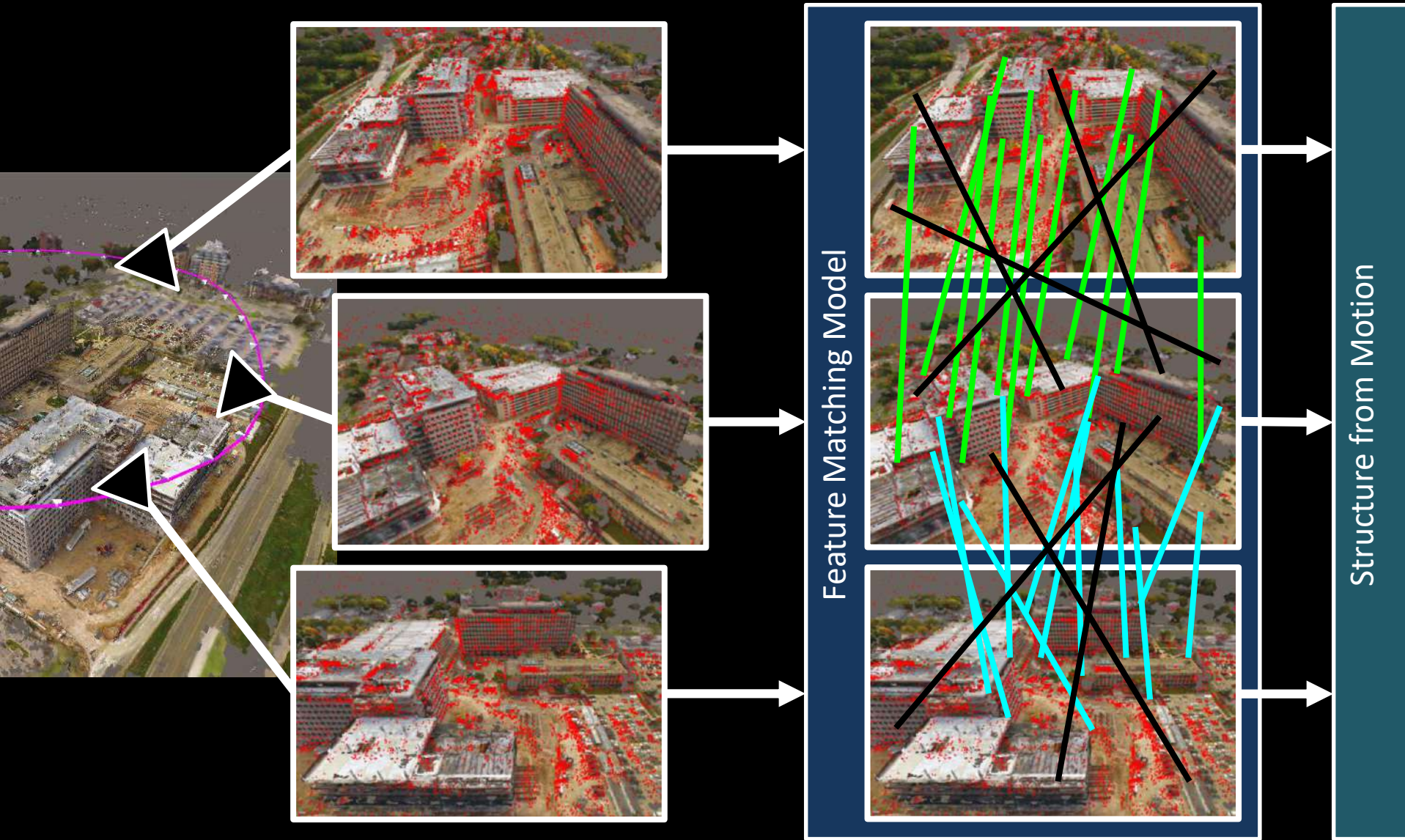
Simulating the Data Capture

Simulates Matching the Synthetic Features



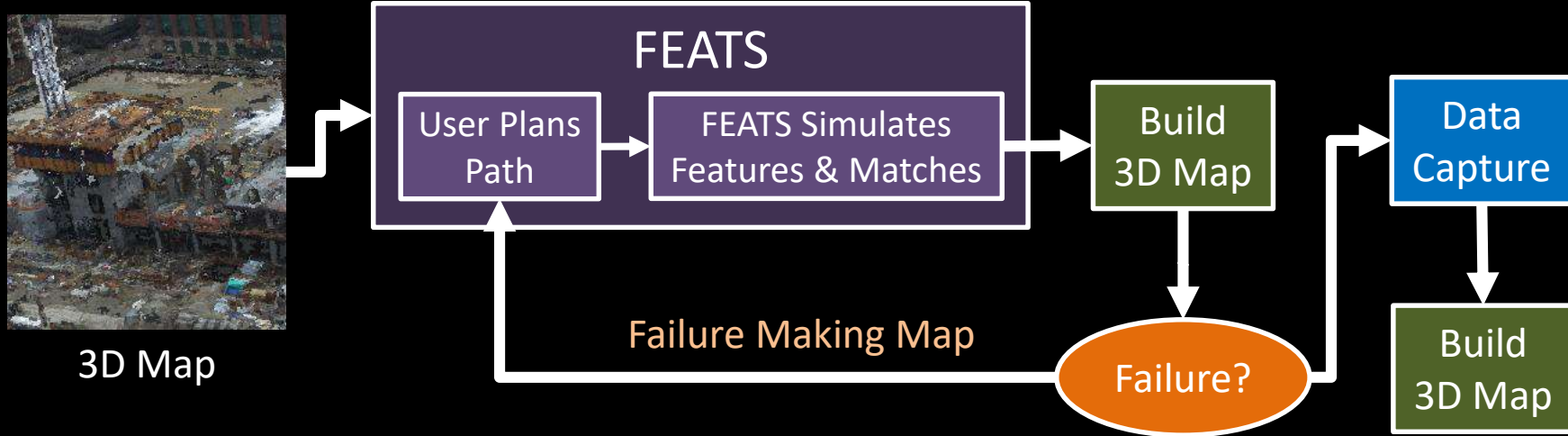
Simulating the Data Capture

Input Features and Matches (tracks) are ready to input to SfM



Example Workflow for FEATS

Week (N)

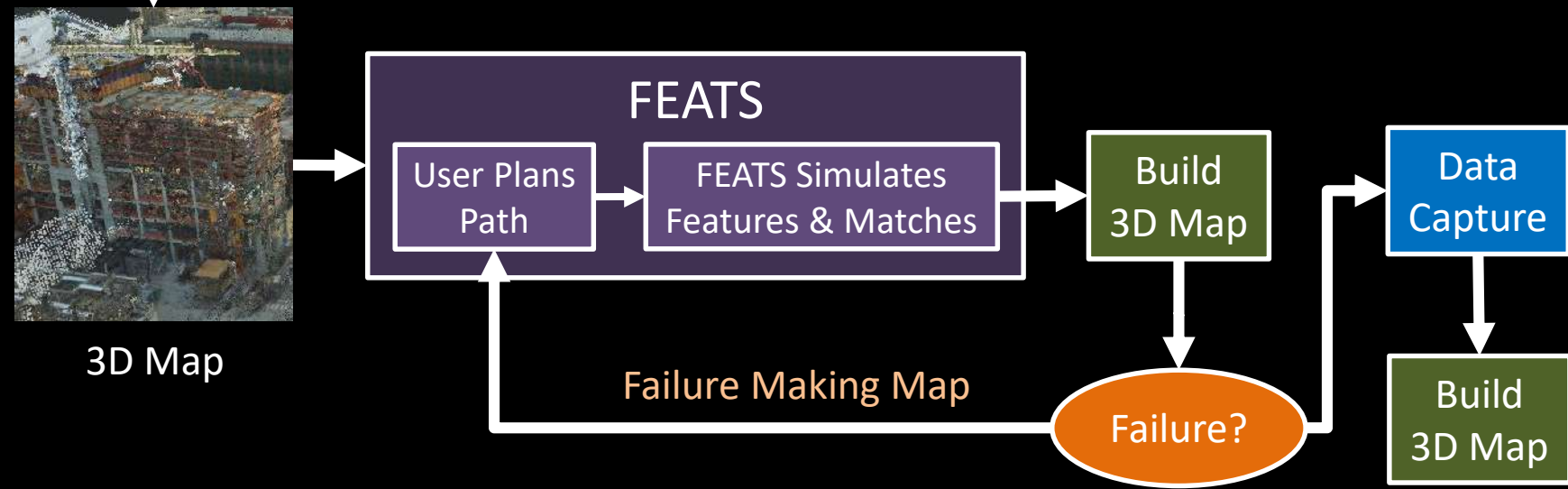


Example Workflow for FEATS

Week (N)



Week (N+1)



FEATS: Synthetic Feature Tracks for Structure from Motion Evaluation

Modeling Feature Noise

Modeling Feature Noise

For All Camera Pairs:

For All 3D Points Seen by Both Cameras:

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$

If $\text{rand}() < P_{\text{Match}}$:

Add Noise for 2D Observations

Add Match for 3D Point

Add Bad Matches for Camera Pair

Modeling Feature Noise

For All Camera Pairs:

For All 3D Points Seen by Both Cameras:

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$

If $\text{rand}() < P_{\text{Match}}$:

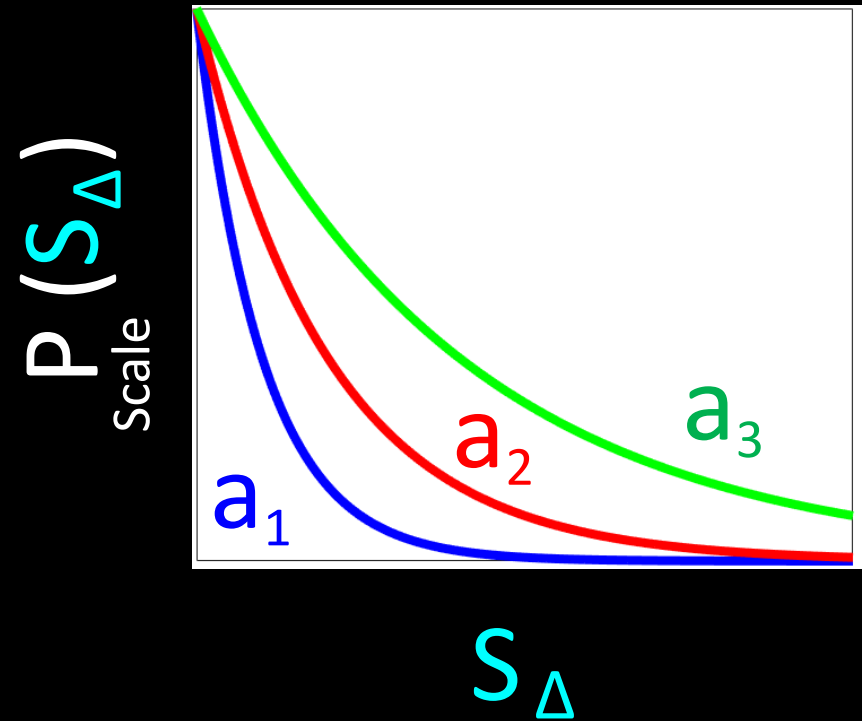
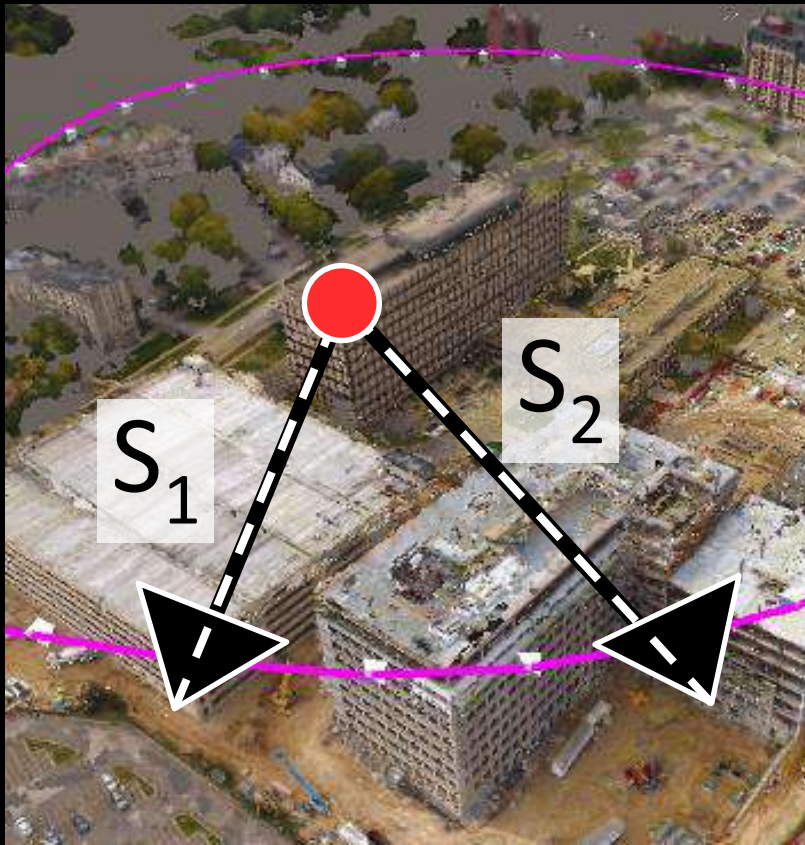
Add Noise for 2D Observations

Add Match for 3D Point

Add Bad Matches for Camera Pair

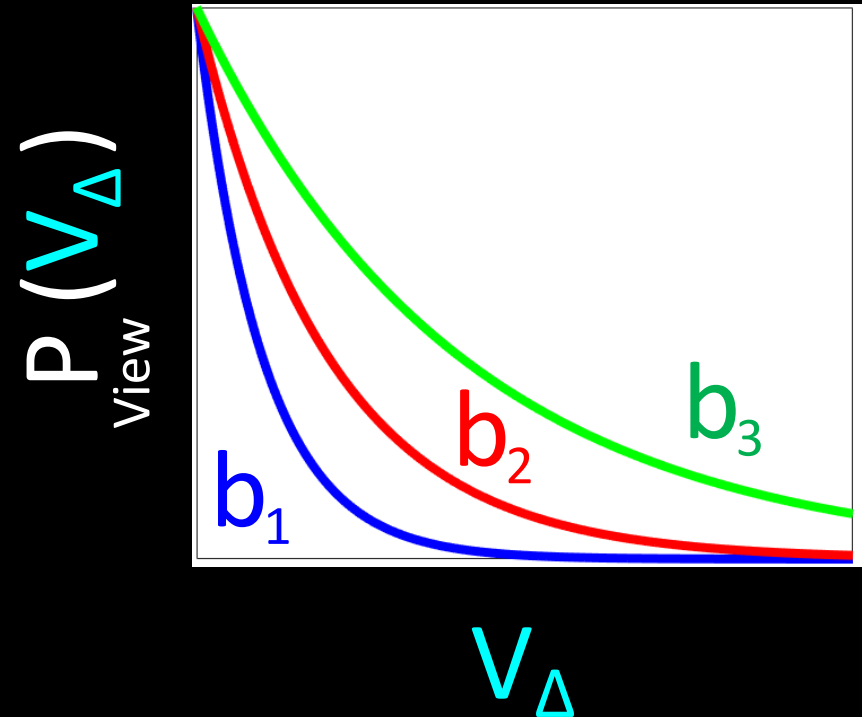
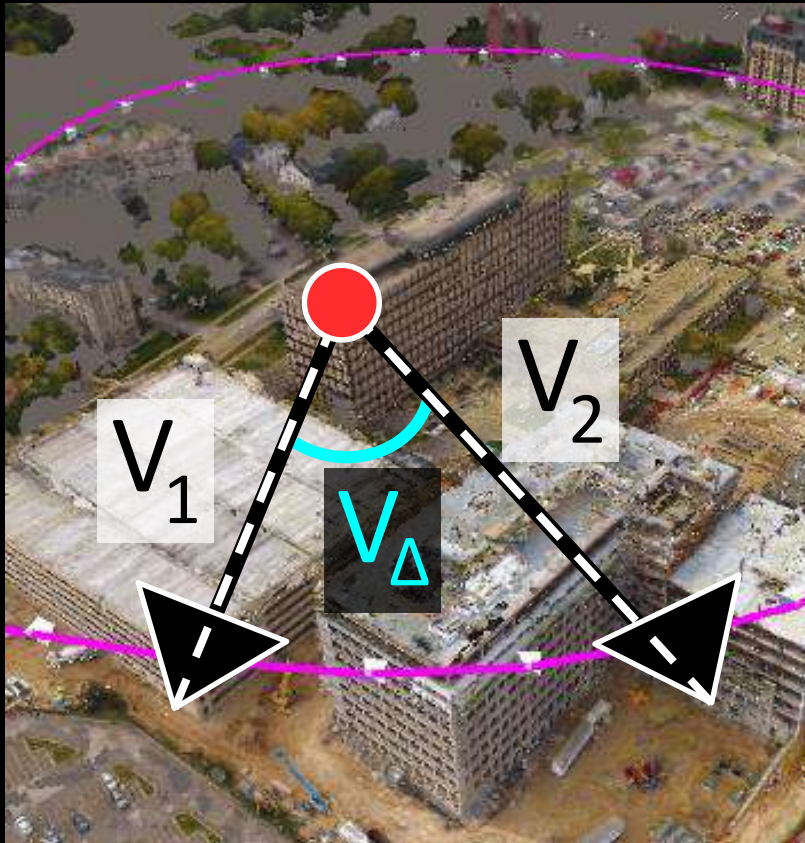
Differences in Scale Decrease Matching Probability

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$



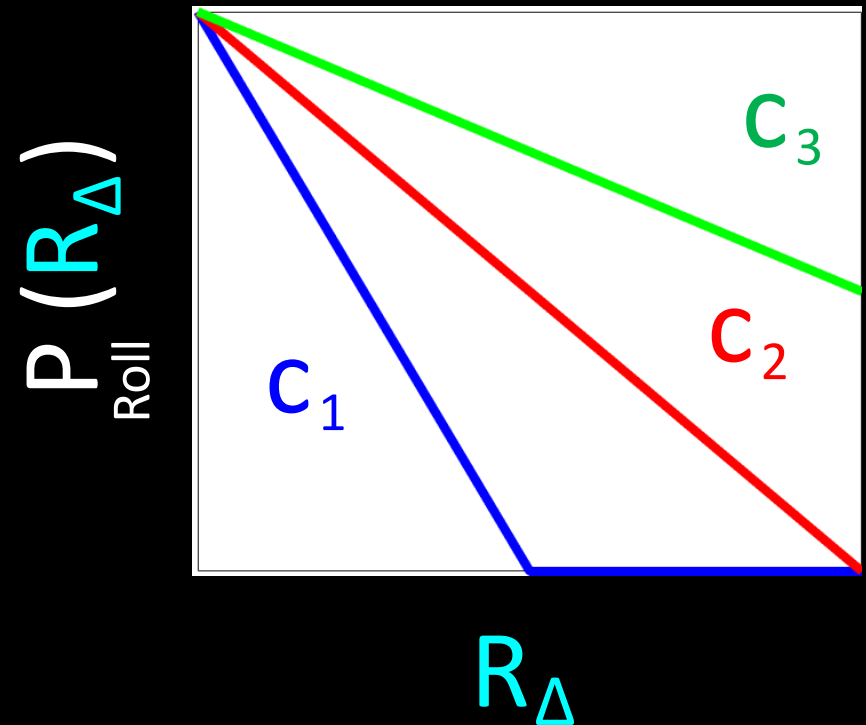
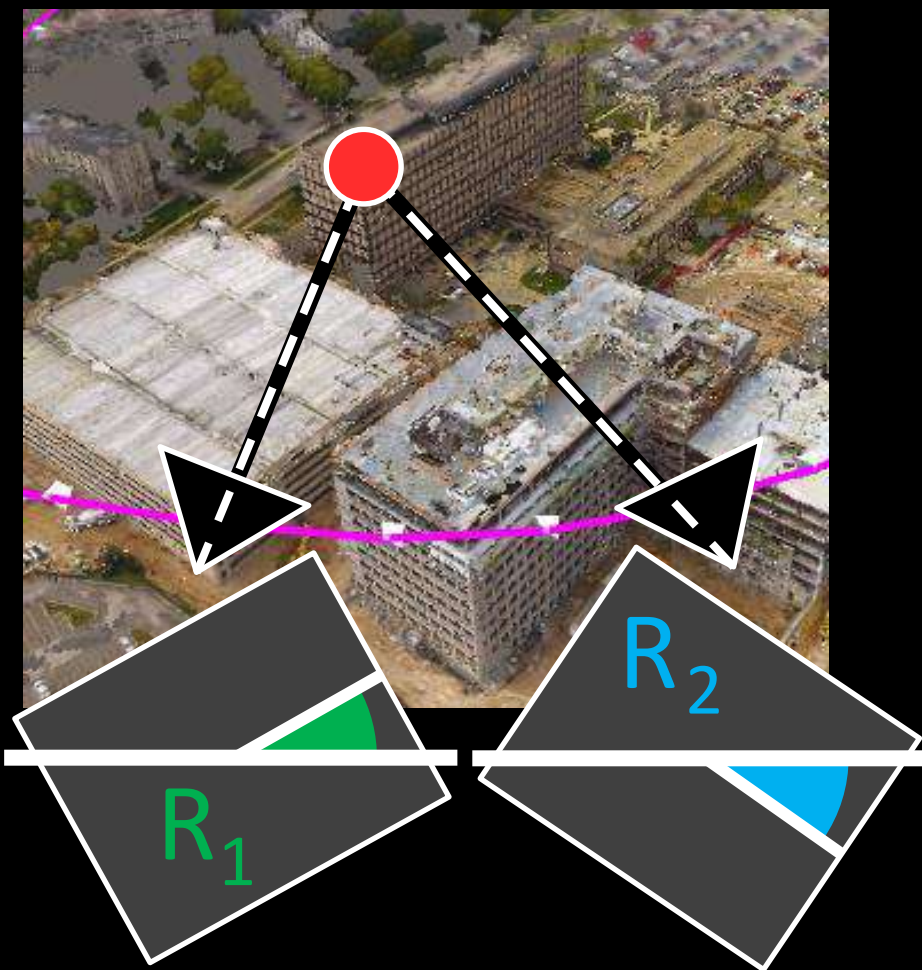
Differences in View Decrease Matching Probability

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$



Differences in Roll Decrease Matching Probability

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$



Modeling Feature Noise

For All Camera Pairs:

For All 3D Points Seen by Both Cameras:

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$

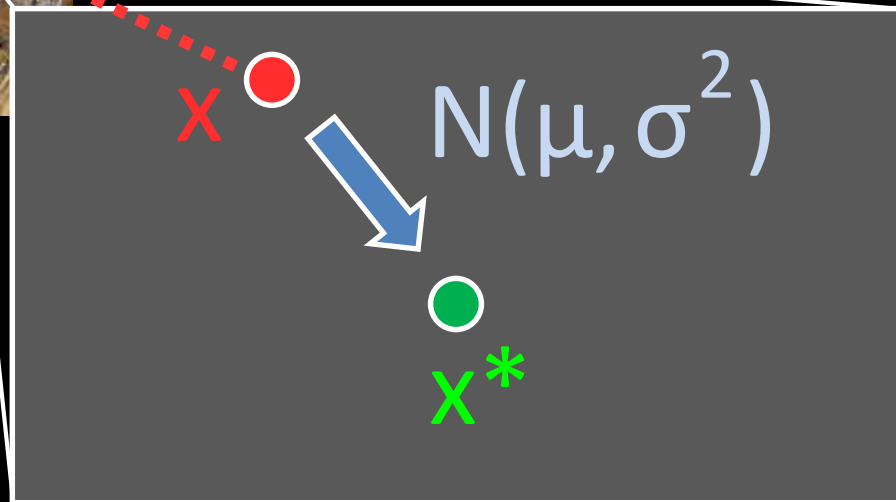
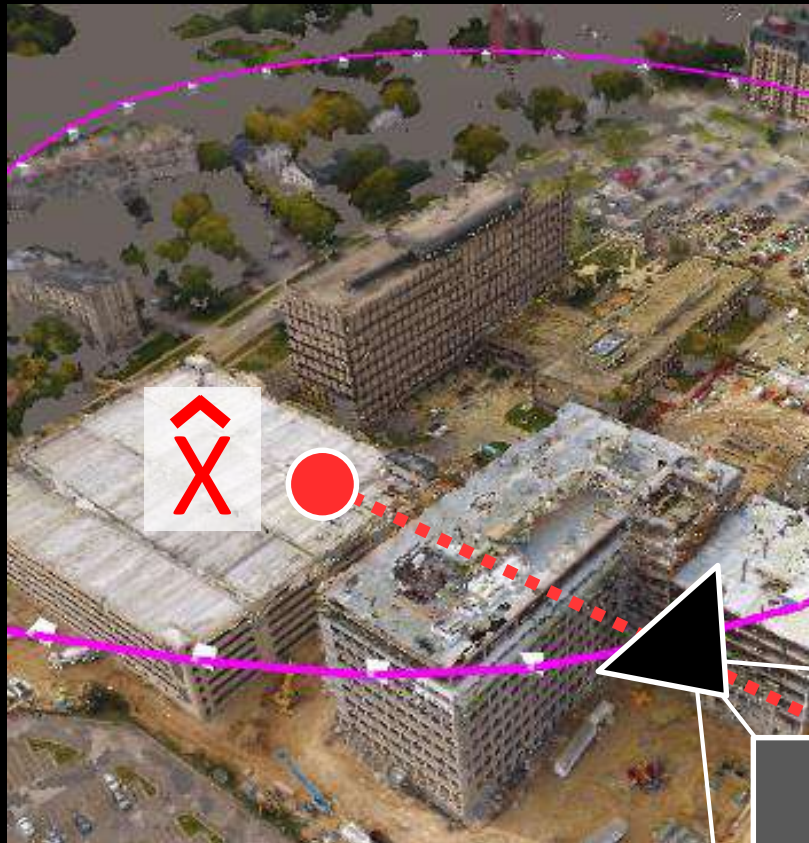
If $\text{rand}() < P_{\text{Match}}$:

Add Noise for 2D Observations

Add Match for 3D Point

Add Bad Matches for Camera Pair

Add Noise for 2D Observations



Modeling Feature Noise

For All Camera Pairs:

For All 3D Points Seen by Both Cameras:

$$P_{\text{Match}} = P_{\text{Scale}}(S_{\Delta}) * P_{\text{View}}(V_{\Delta}) * P_{\text{Roll}}(R_{\Delta})$$

If $\text{rand}() < P_{\text{Match}}$:

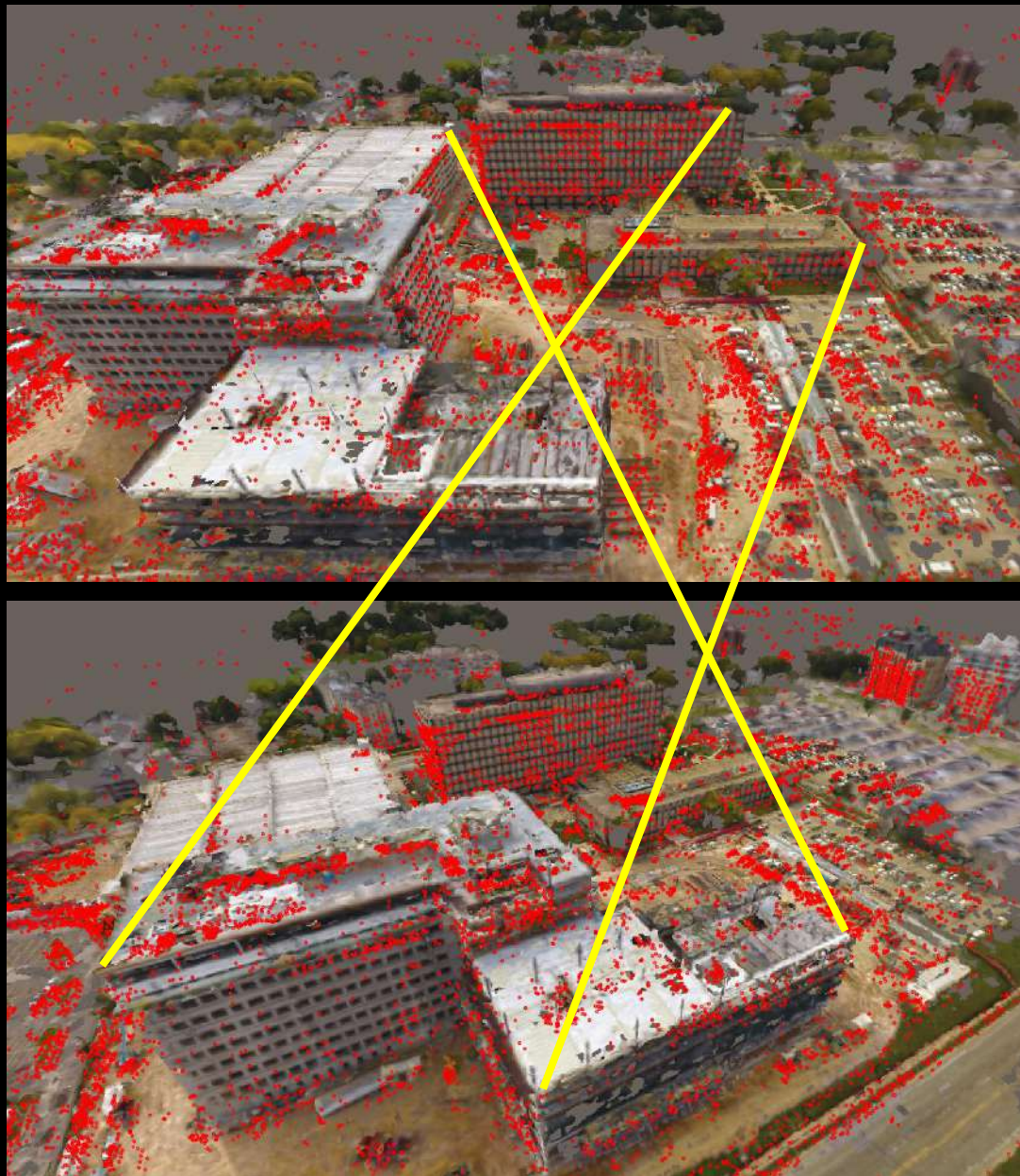
Add Noise for 2D Observations

Add Match for 3D Point

Add Bad Matches for Camera Pair

Feature Match Model: Bad Matches

Add bad matches

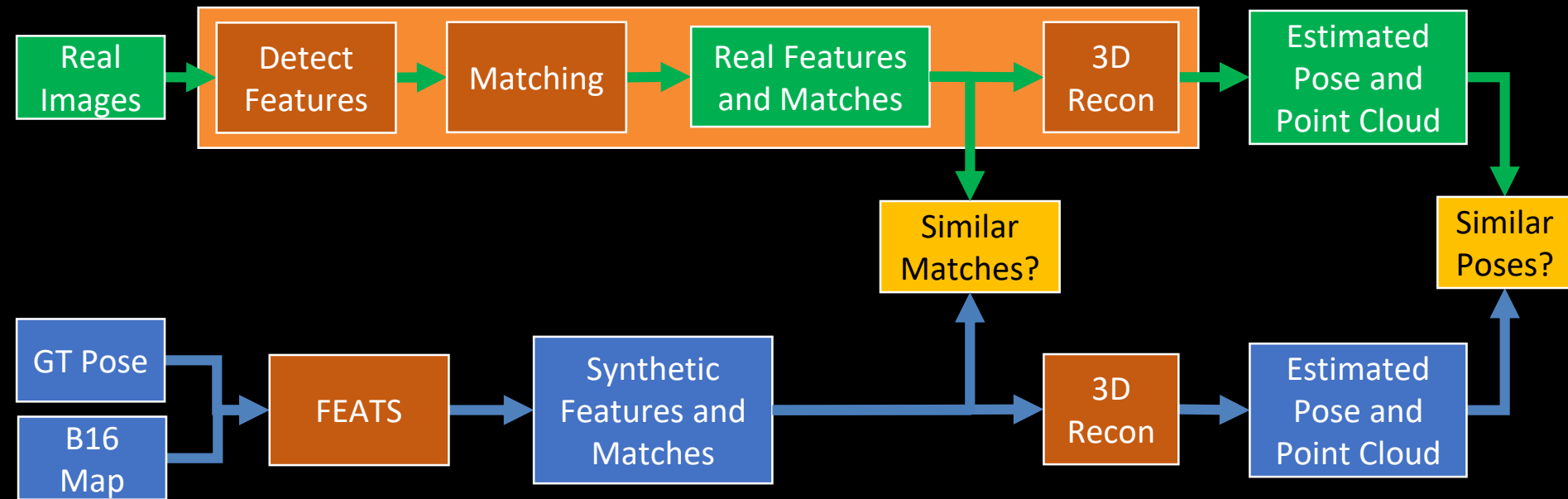


FEATS: Synthetic Feature Tracks for Structure from Motion Evaluation

Comparing to Real Data

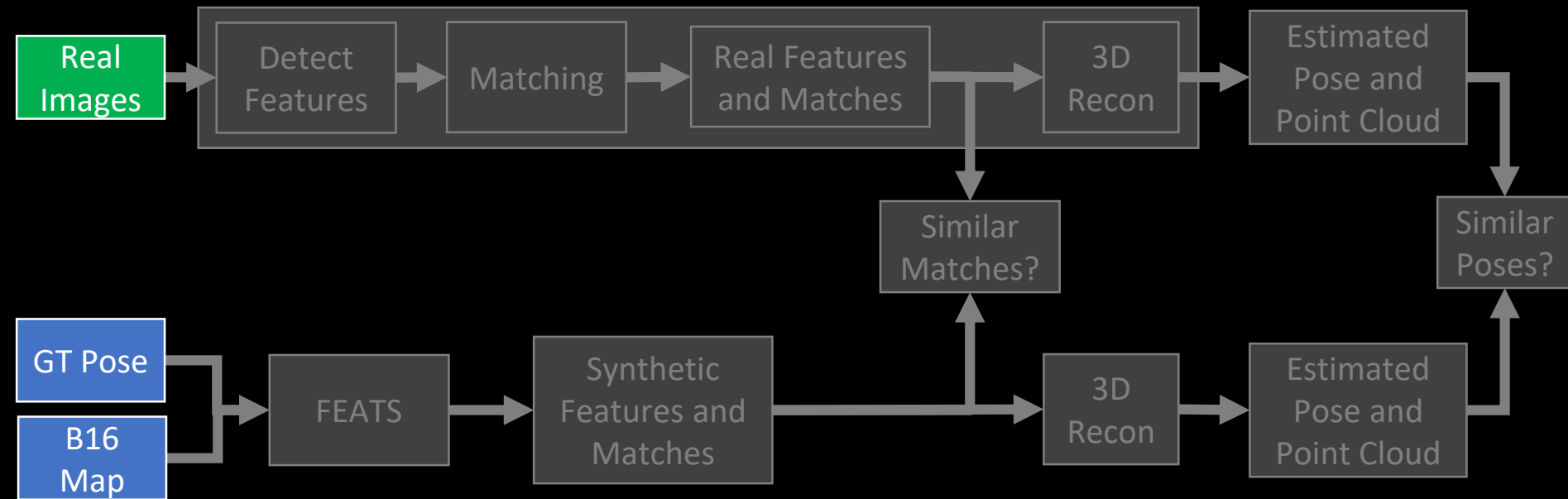
Validating the Simulator

Show that the simulator provides results representative of the real world.

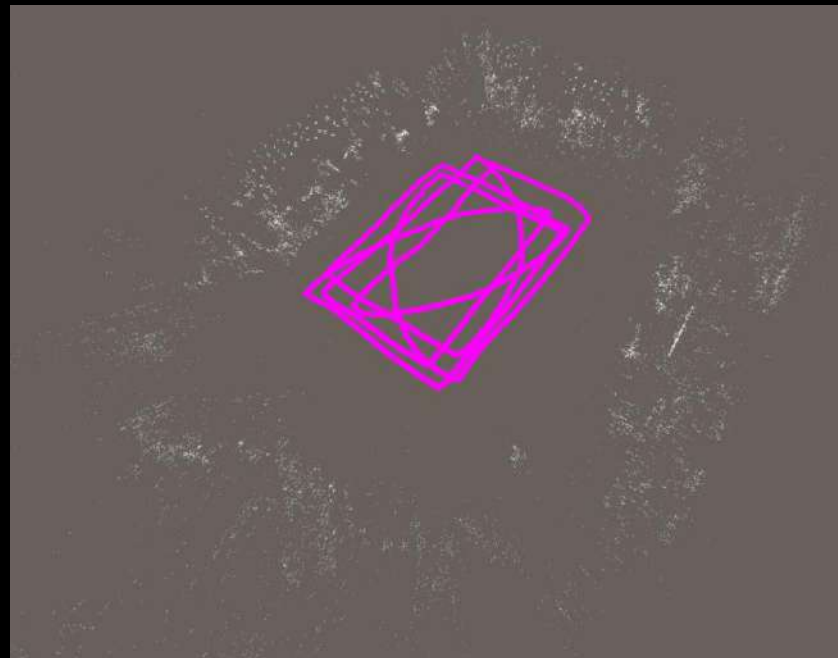


Validating the Simulator

Show that the simulator provides results representative of the real world.



Mapping B16



16 Validation Trajectories



Arc 1



Arc 2



Arc 3



Egg



Long 1



Long 2



Long 3



Long 4



Long 5



Rotation Fast



Rotation Slow



Snake 1



Snake 2



Straight 1



Straight 2

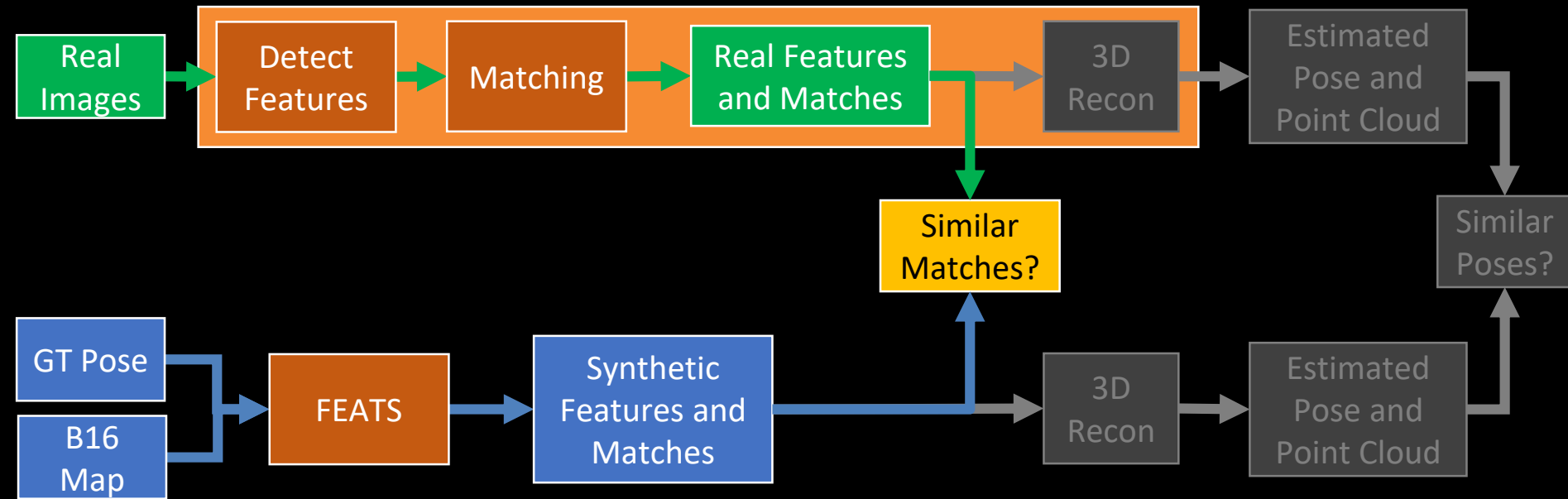


X

Images and Ground Truth Pose for each

Validating the Simulator

Show that the simulator provides results representative of the real world.



COLMAP
CVPR & ECCV 2016



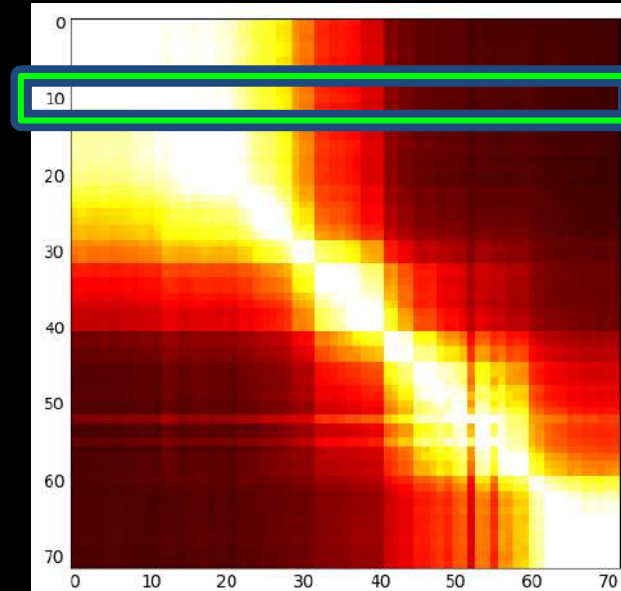
OpenSfM
by Mappilary



VisualSfM
3DV 2013

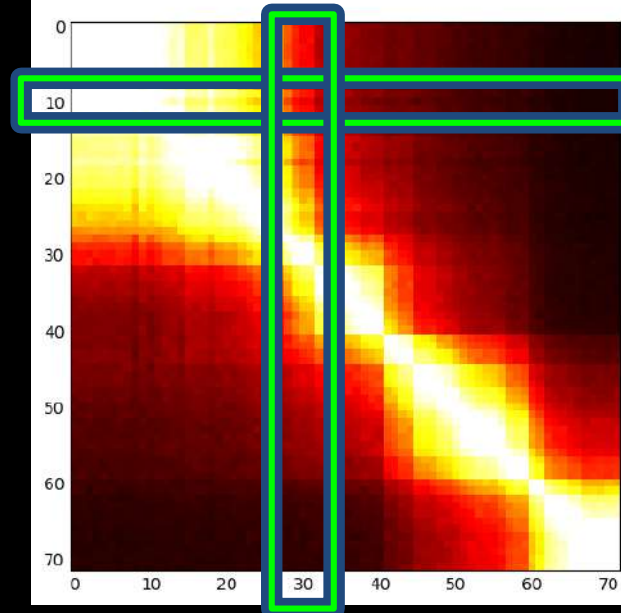
Verifying the Match Model

Real



Percent feature matches between image 10 and images 0 to 70

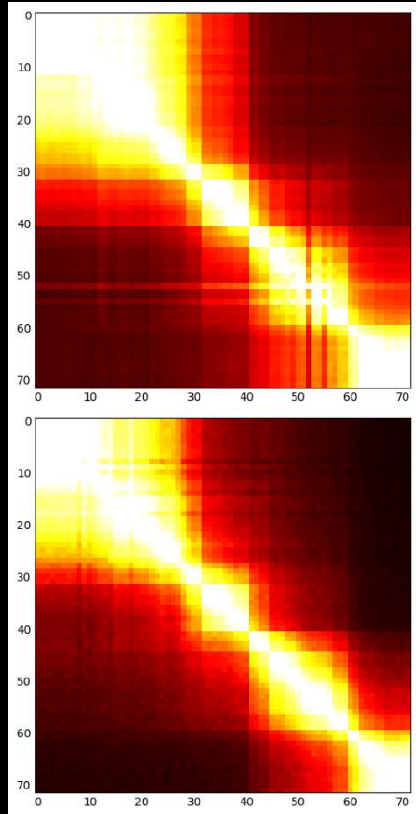
Synthetic



Percent feature matches between image 10 and image 30

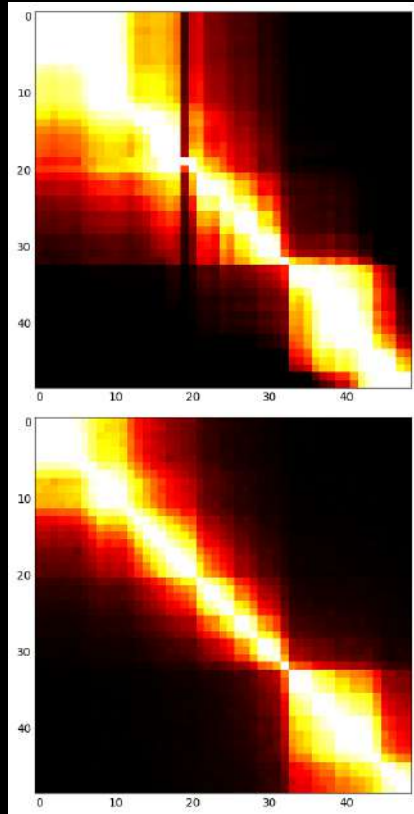
Verifying the Match Model

ALL correlations above +0.74



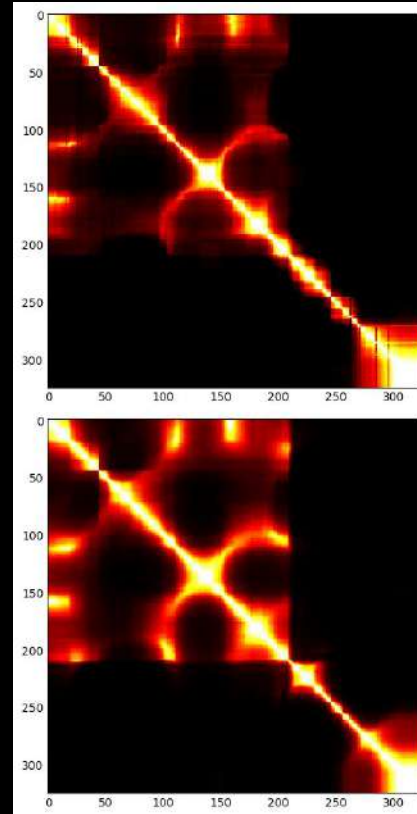
Straight 2 (OpenSfM)

0.98



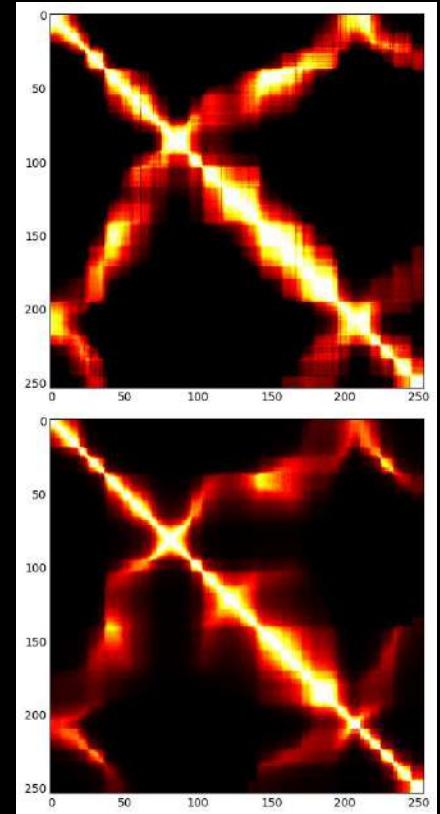
Rot. Fast (COLMAP)

0.94



Long 3 (OpenSfM)

0.93

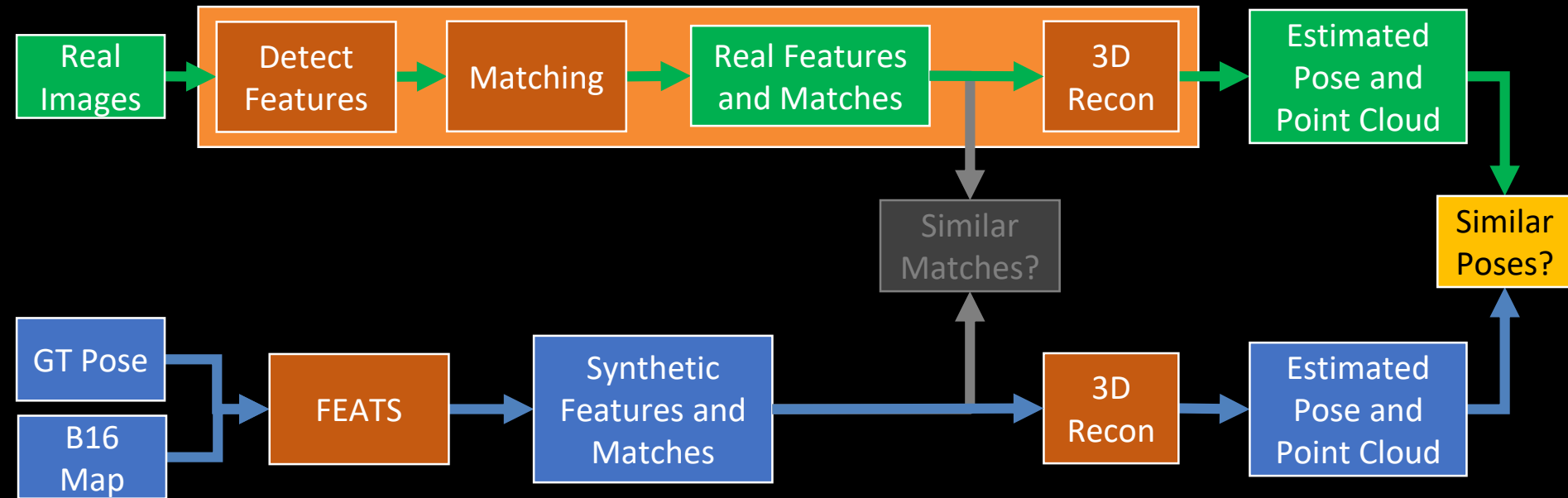


Long 5 (VisualSfM)

0.88

Validating the Simulator

Show that the simulator provides results representative of the real world.



COLMAP
CVPR & ECCV 2016



OpenSfM
by Mappilary



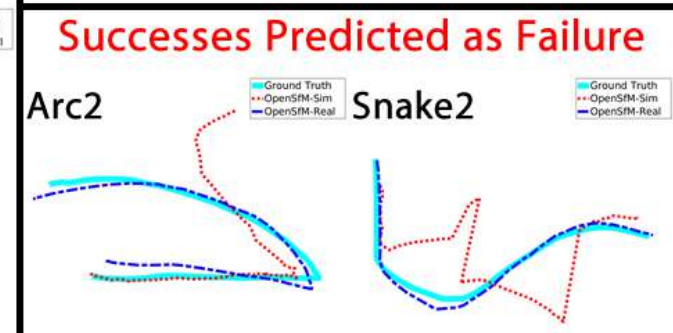
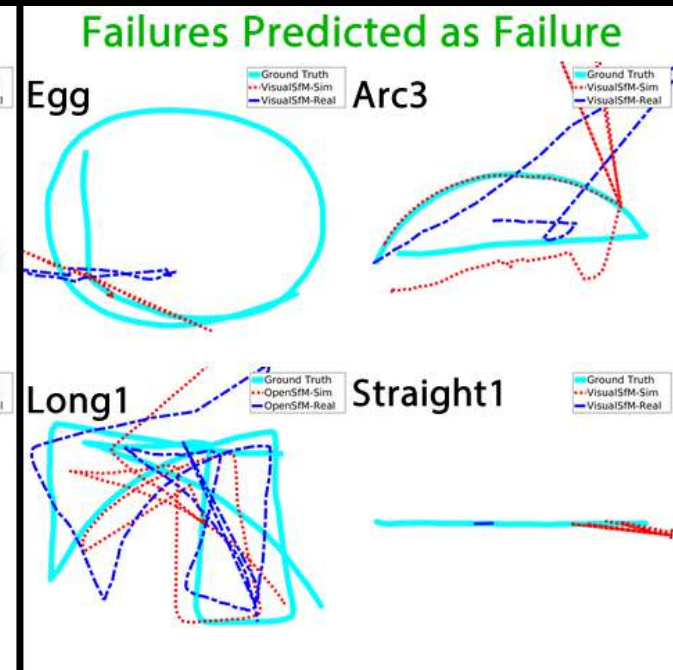
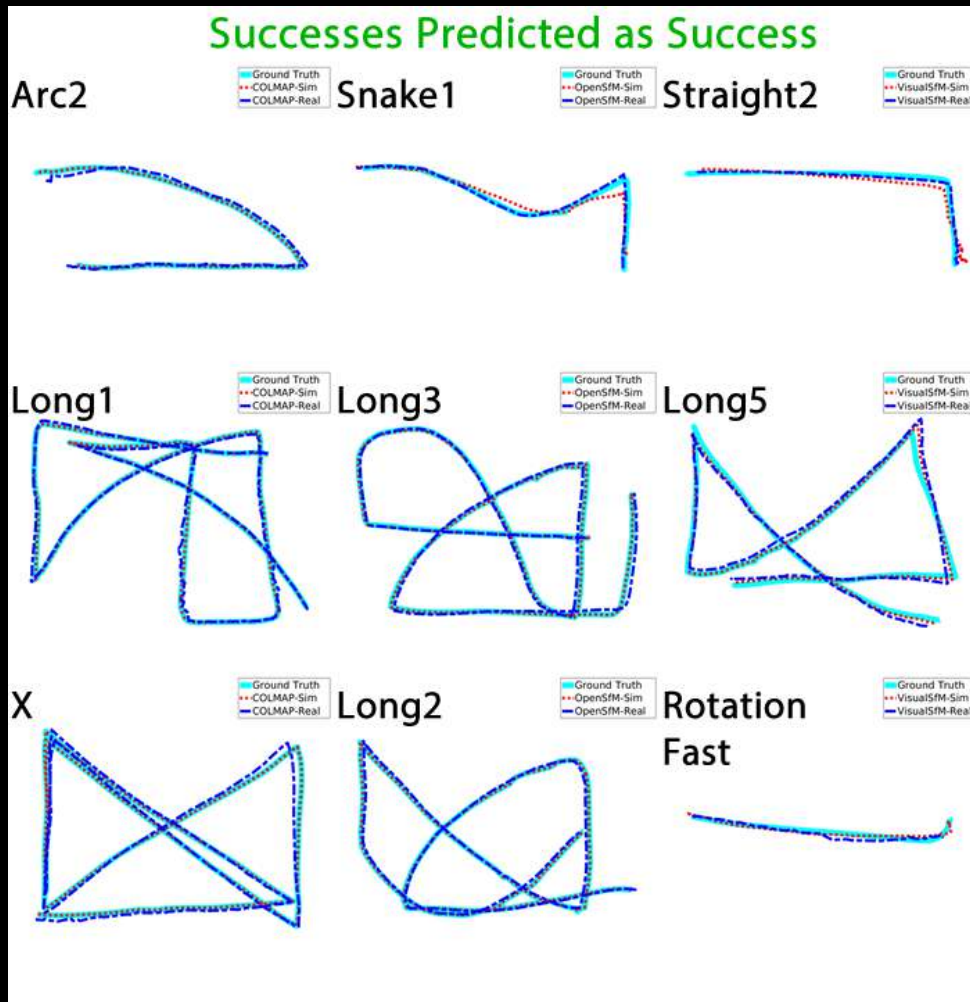
VisualSfM
3DV 2013

Verifying 48 Synthetic Reconstructions

38 Successes Predicted as Success

8 Failures Predicted as Failure

2 Successes Predicted as Failure

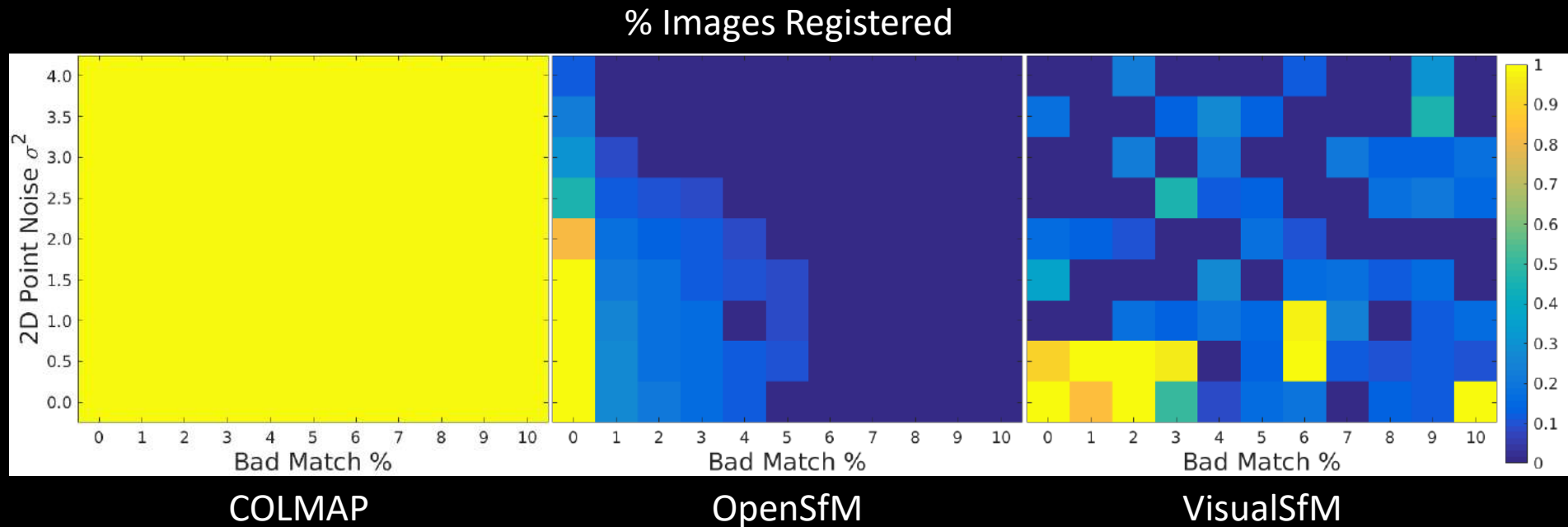


FEATS: Synthetic Feature Tracks for Structure from Motion Evaluation

New Evaluations

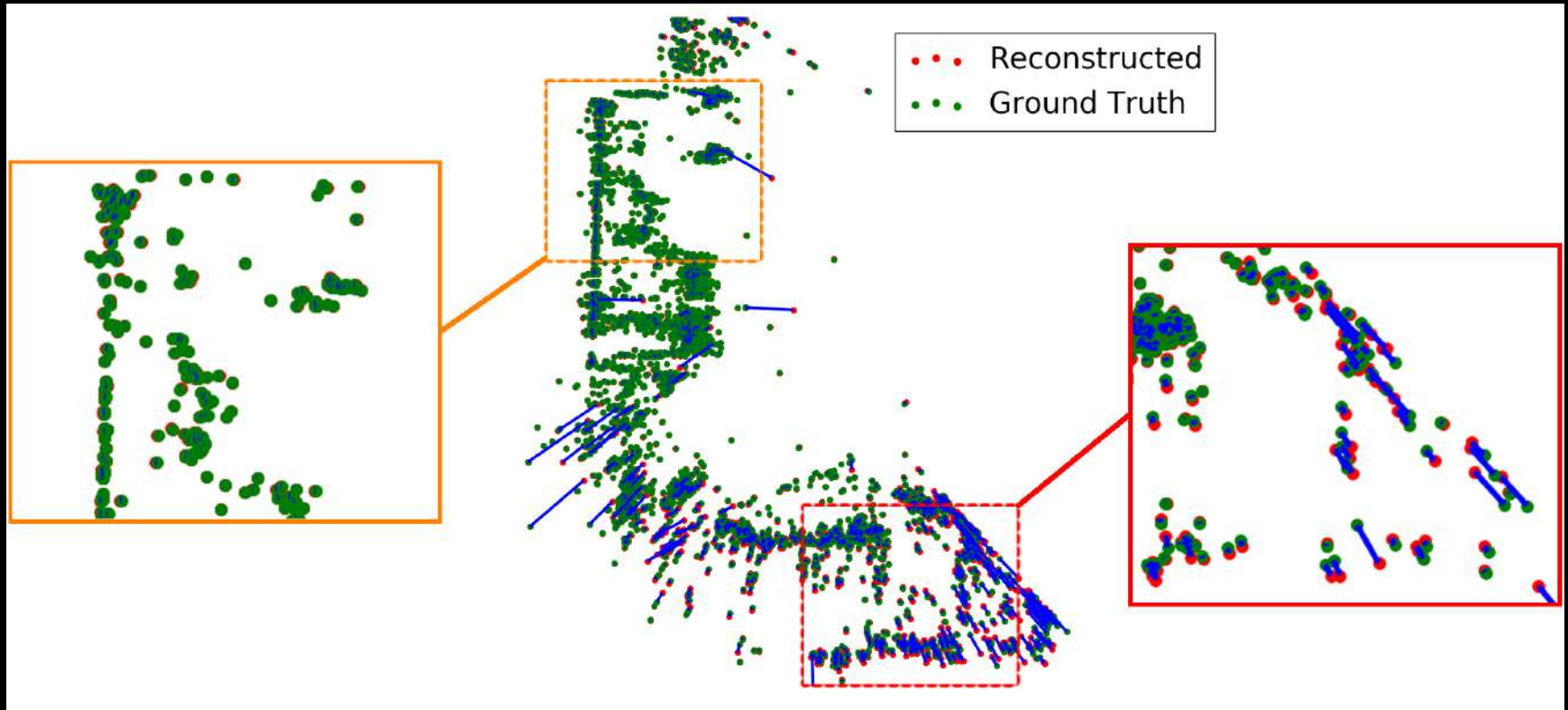
Varying Feature Noise and Bad Matches

FEATS provides control of noise parameters and hundreds of different trajectories can be generated quickly



3D Point Error

FEATS provides ground truth 3D location for each estimated 3D point

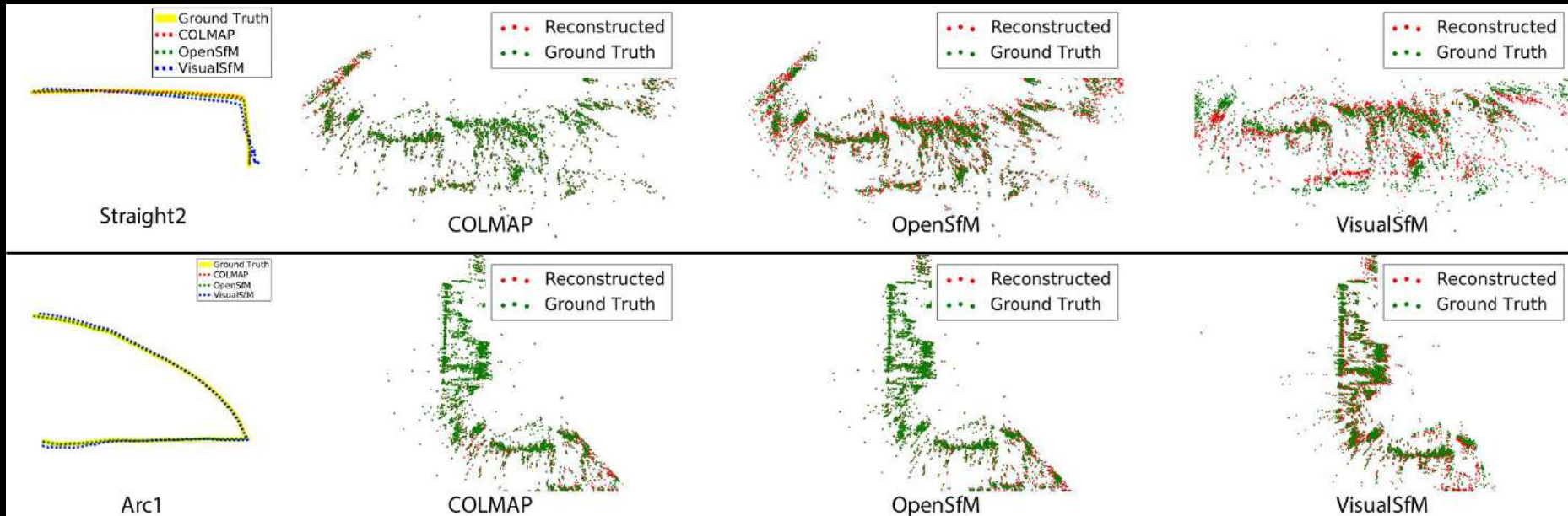


Traditional 3D geometry ground truth methods (Laser Scanners) do not

3D Point Error

Similar trajectory errors

Different geometry errors



	COLMAP	OpenSfM	VisualSfM
Arc 1	21.5 mm	26.0 mm	84.7 mm
Straight 2	27.2 mm	118.1 mm	246.8 mm

Metrics to evaluate 3D point error are important

Vision Based Robots Monitoring Built Environments

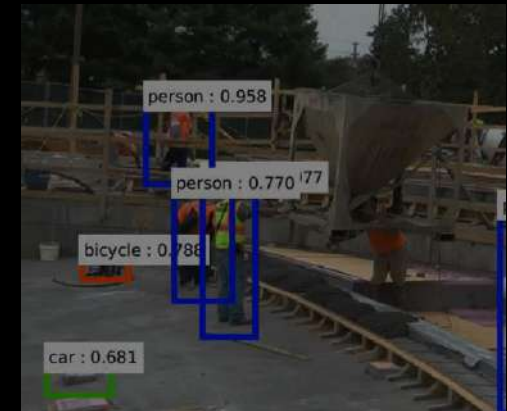
Data Capture



Mapping

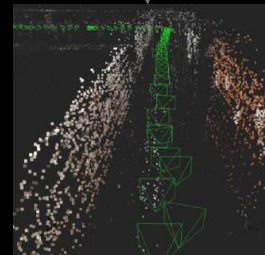


Analysis



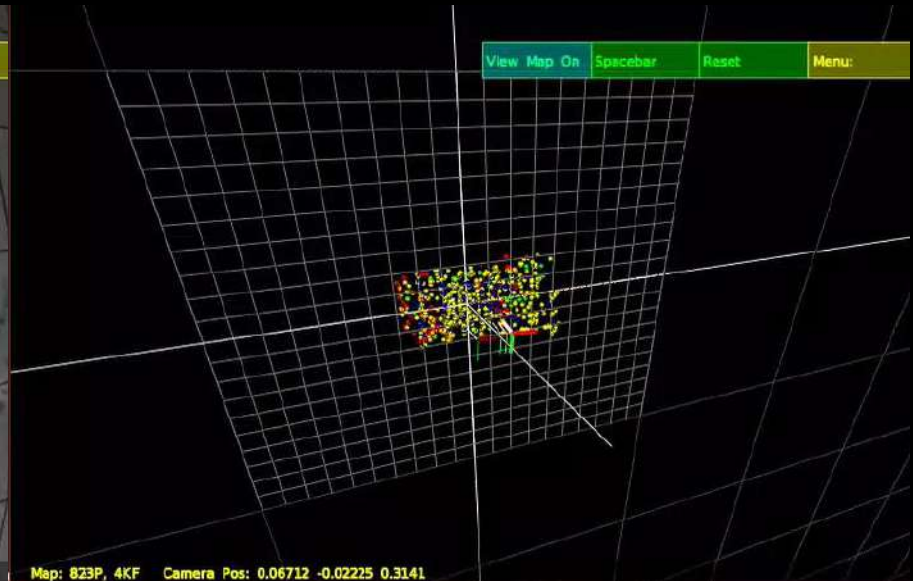
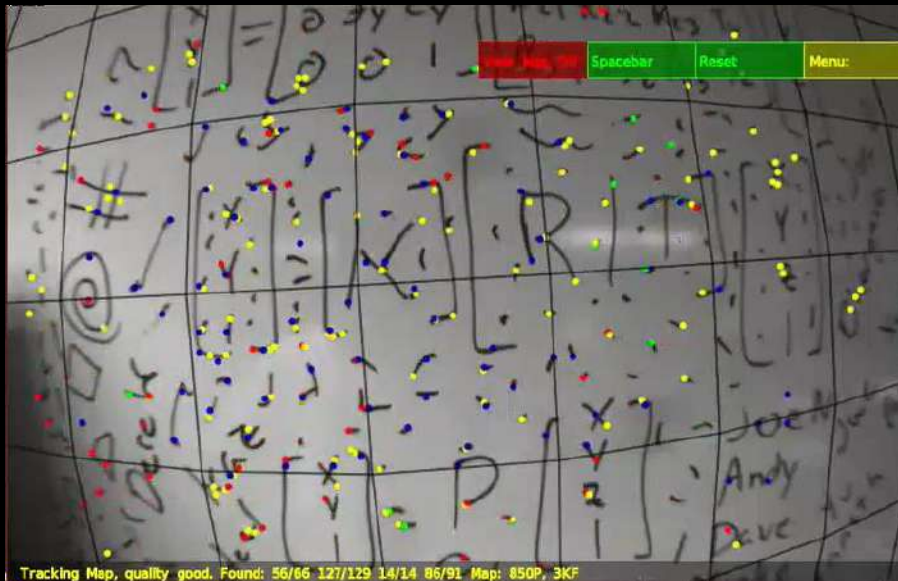
SLAM

Structure from Motion



Simultaneous Localization And Mapping

- SLAM also does 3D Reconstruction
- SLAM runs in real time on video
- SfM runs offline on unordered images



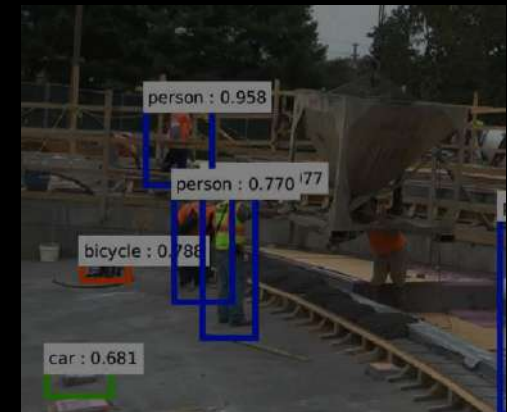
SLAM Fails When Few Features Are Trackable

Vision Based Robots Monitoring Built Environments

Data Capture

Mapping

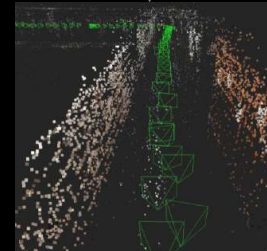
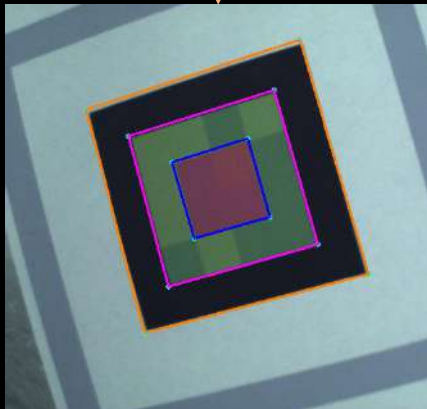
Analysis



SLAM

Structure from Motion

Overcoming
Failures



A Colored Marker and Fast
Detection Algorithm

ChromaTag: A Colored Fiducial Marker and Fast Detection Algorithm

*Joseph DeGol,
Timothy Bretl,
Derek Hoiem*

ICCV 2017

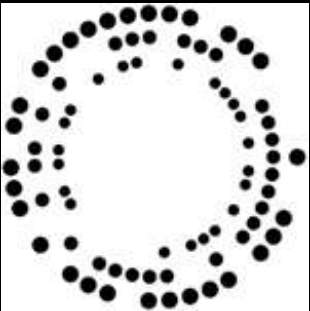
ChromaTag: A Marker for SLAM

Other markers are too slow for SLAM

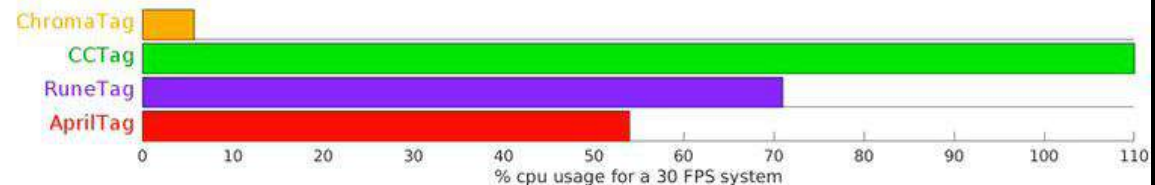
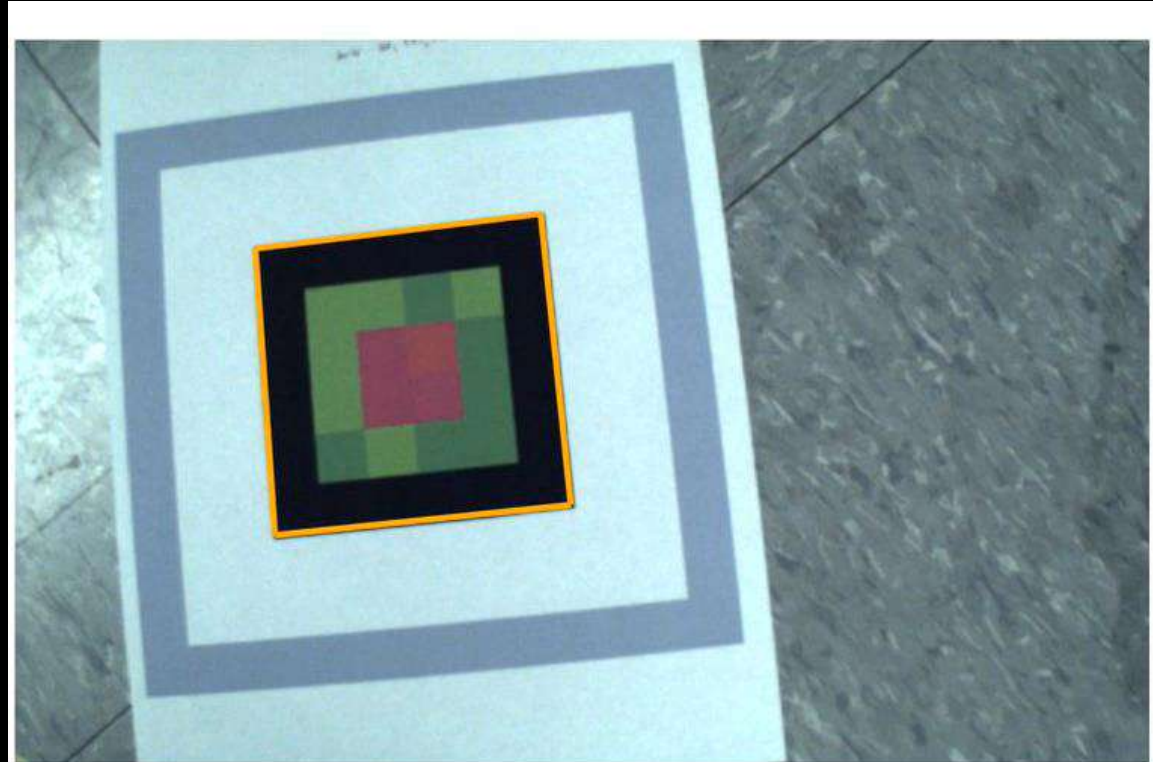
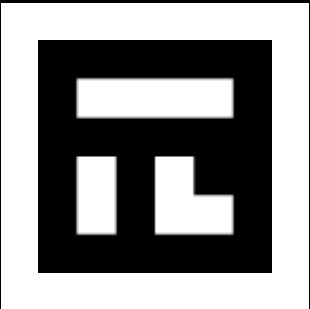
CCTag



RuneTag



AprilTag



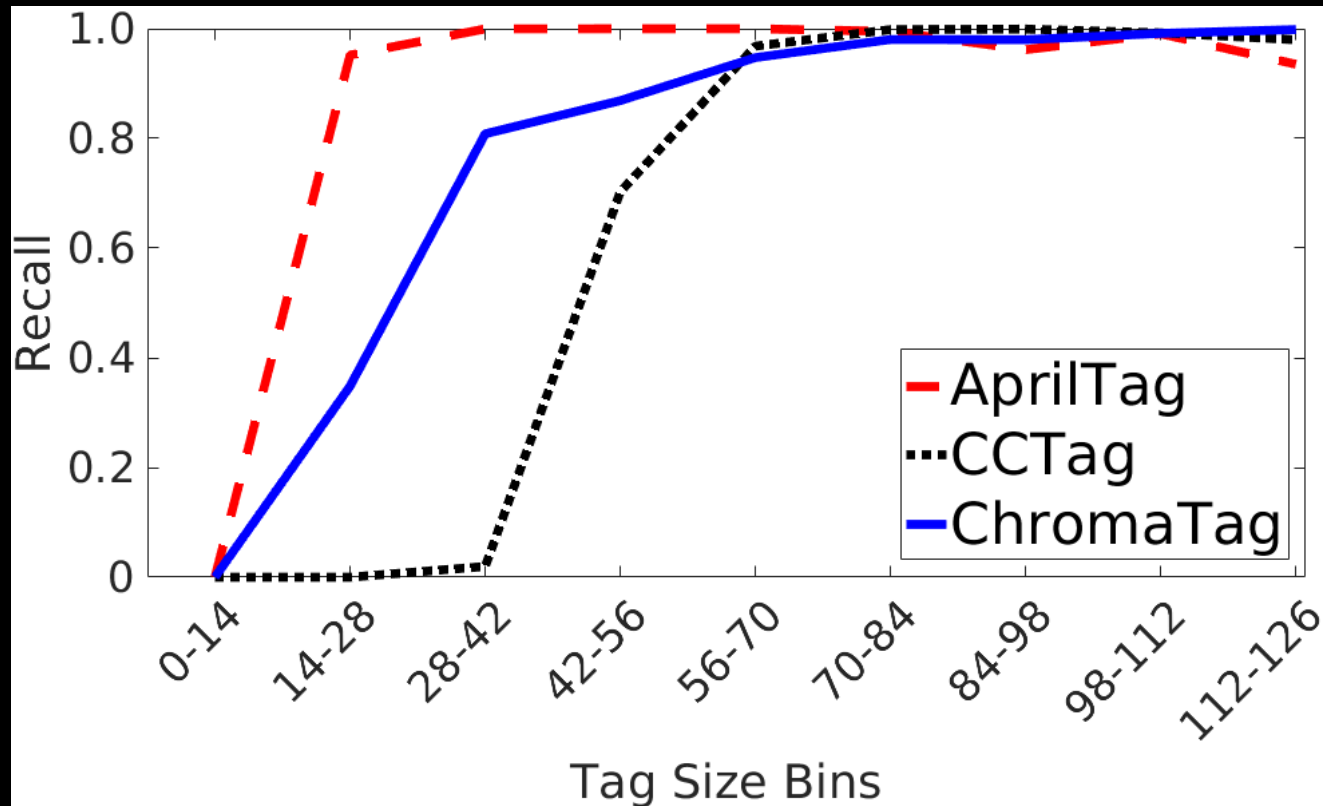
Average Frames Per Second

ChromaTag is significantly faster than AprilTag, CCTag, and RuneTag for both >0 and 0 detections.

	Average Frames Per Second		
	Total	> 0 Detections	0 Detections
ChromaTag	926	709	2616
AprilTag	56	56	49
CCTag	10	7	19
RuneTag	42	2	71

Recall vs. Tag Size and Viewing Angle

Recall > 80% after 30*30 pixel tag size



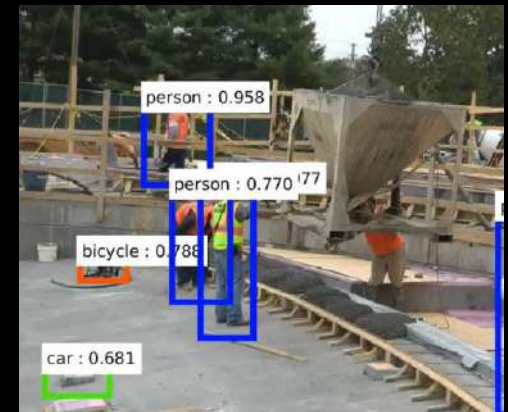
Precision is ~96%

Vision Based Robots Monitoring Built Environments

Data Capture

Mapping

Analysis



SLAM

Structure from Motion

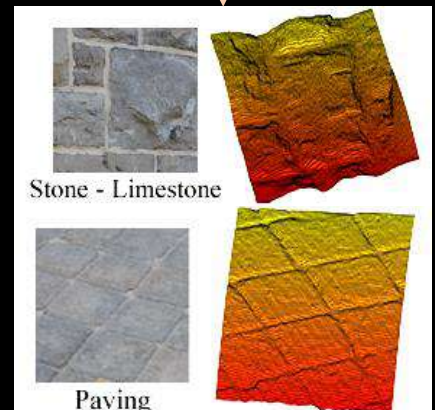
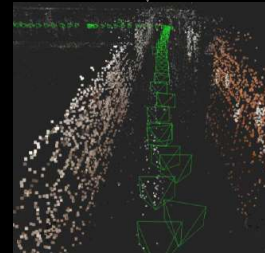
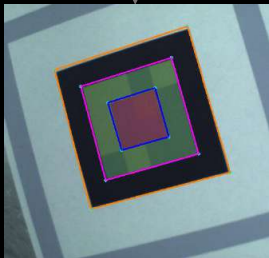


Image and Geometry for
Material Recognition

Geometry Informed Material Recognition

Joseph DeGol, Mani Golparvar-Fard, Derek Hoiem
CVPR 2016 (Spotlight Paper)

Recognize materials in real world scenes using images (2D) and geometry (3D).



Asphalt



Brick



Concrete



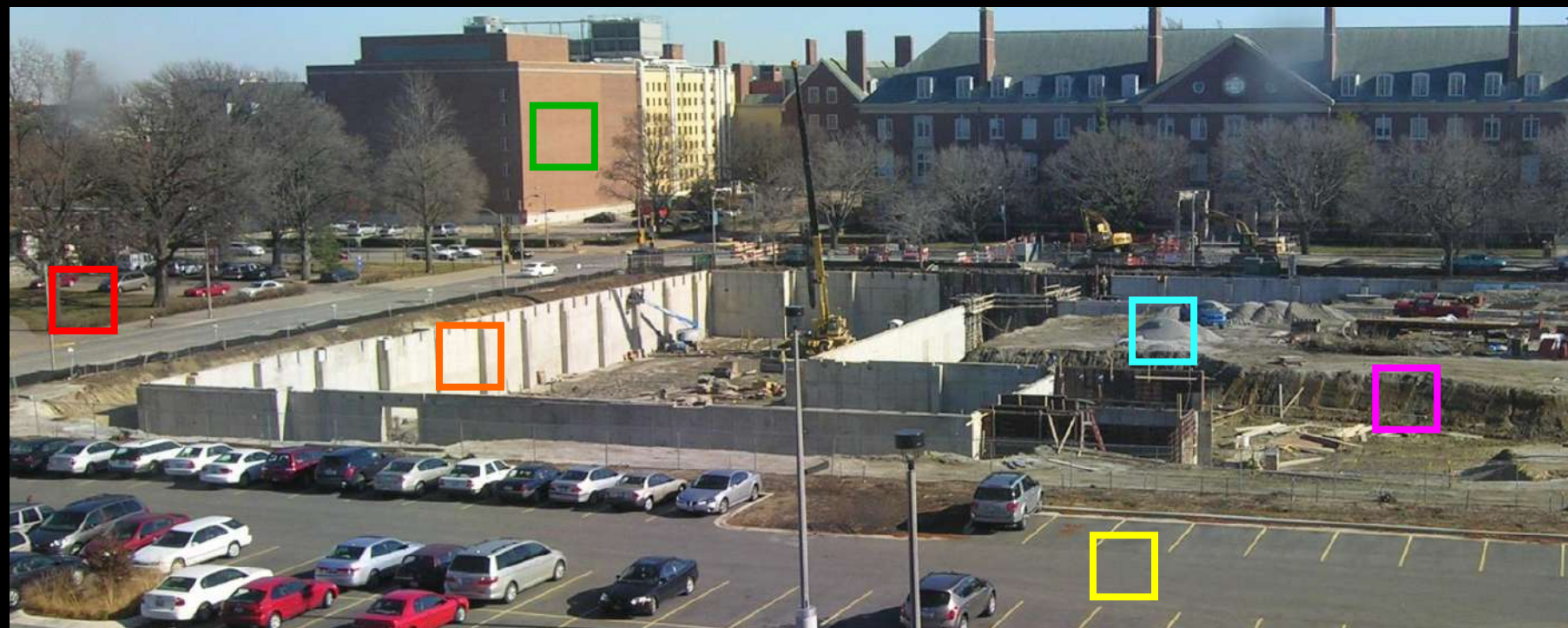
Grass



Gravel



Soil



Challenges

Appearance changes due to lighting



Challenges

Appearance changes due to lighting



Appearance changes due to perspective



Challenges

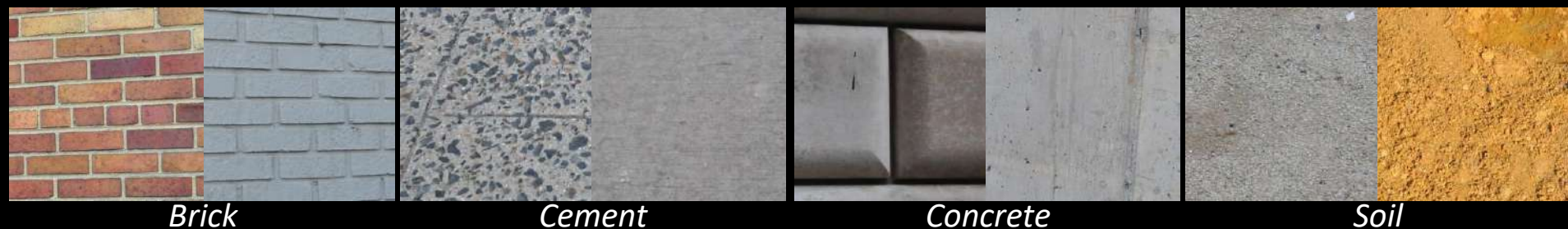
Appearance changes due to lighting



Appearance changes due to perspective



Appearance is similar within and across categories



Brick

Cement

Concrete

Soil

Geometry Aids Material Recognition

3D Geometry helps with categories with similar appearance but different geometry.

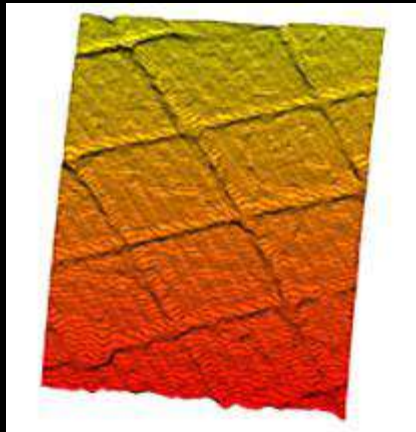
Paving



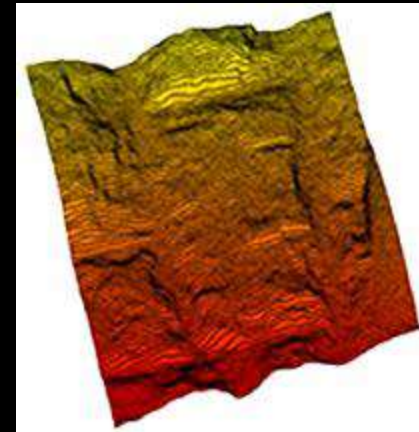
Stone - Limestone



Often confused with 2D



Correctly classified with 3D



Best 2D (FV+CNN)

68.92

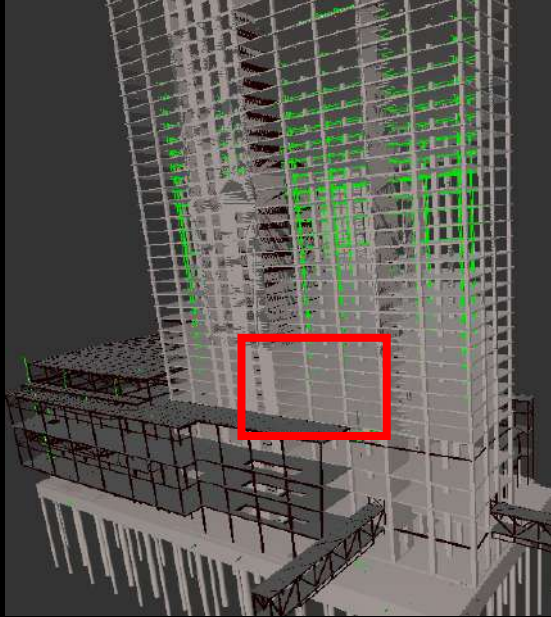
Best 3D (FV-N+CNN+N3D)

73.84

Geometry and Appearance Based Reasoning of Construction Progress Monitoring

Kevin Han, Joseph DeGol, Mani Golparvar-Fard
ASCE Journal of Construction Engineering and Management

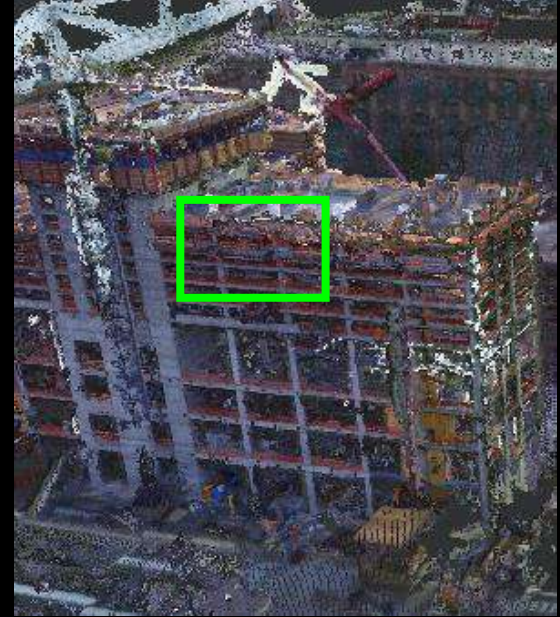
BIM



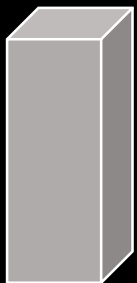
Week 24



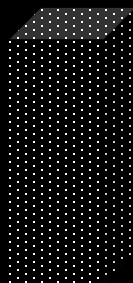
Week 29



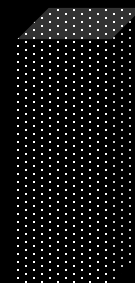
Cement Column



Column : In Progress
Wood

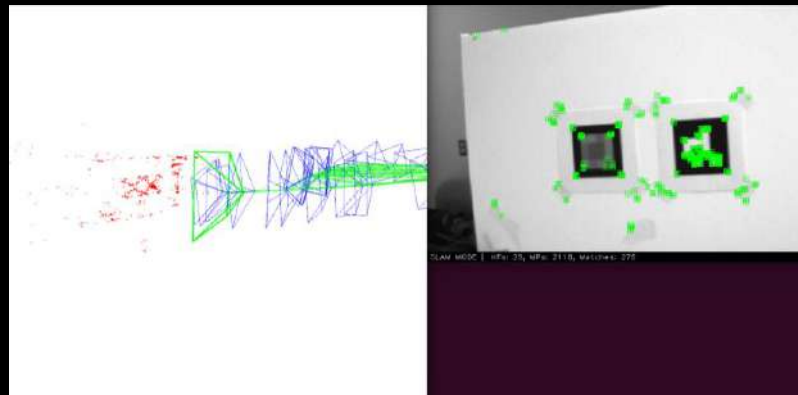
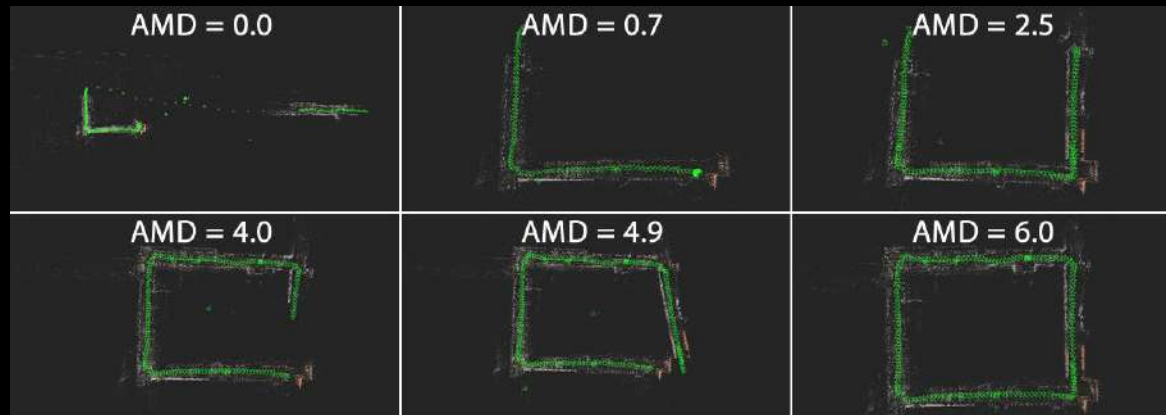


Column : Complete
Cement



Conclusions / Future Work

- Placing the minimum number of markers that still results in successful 3D reconstructions
- Improving Matching and Resectioning without Markers
- Planning paths to improve 3D reconstruction
- Improved SLAM using Markers



Additional Work

Automatic Grasp Selection using a Camera in a Hand Prosthesis

Joseph DeGol, Aadeel Akhtar, Bhargava Manja, Timothy Bretl
EMBC 2016: Best Student Paper Award (3rd Place)

A Passive Mechanism for Relocating Payloads with a Quadrotor

Joseph DeGol, David Hanley, Navid Aghasadeghi, Timothy Bretl
IROS 2015

A Clustering Approach for Detecting Moving Objects Captured by
a Moving Aerial Vehicle

Joseph DeGol, Myra Nam
ICASSP 2014

degol2.web.engr.illinois.edu/

Thanks and Questions



NDSEG Fellowship



3M Fellowship



F30HD084201



CMMI-1427111



CMMI-1544999



IIS-1320519



CMMI-1446765



Derek Hoiem



Tim Bretl



Mani Golparvar



Sudipta Sinha



David Forsyth



Xinke Deng



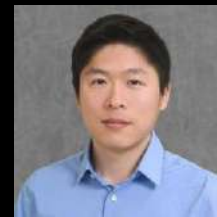
Rajbir Kataria



Bhargava Manja



Aadeel Akhtar



Kevin Han



Jae Yong Lee



Daniel Yuan



Myra Nam



Navid Aghasadeghi



David Hanley