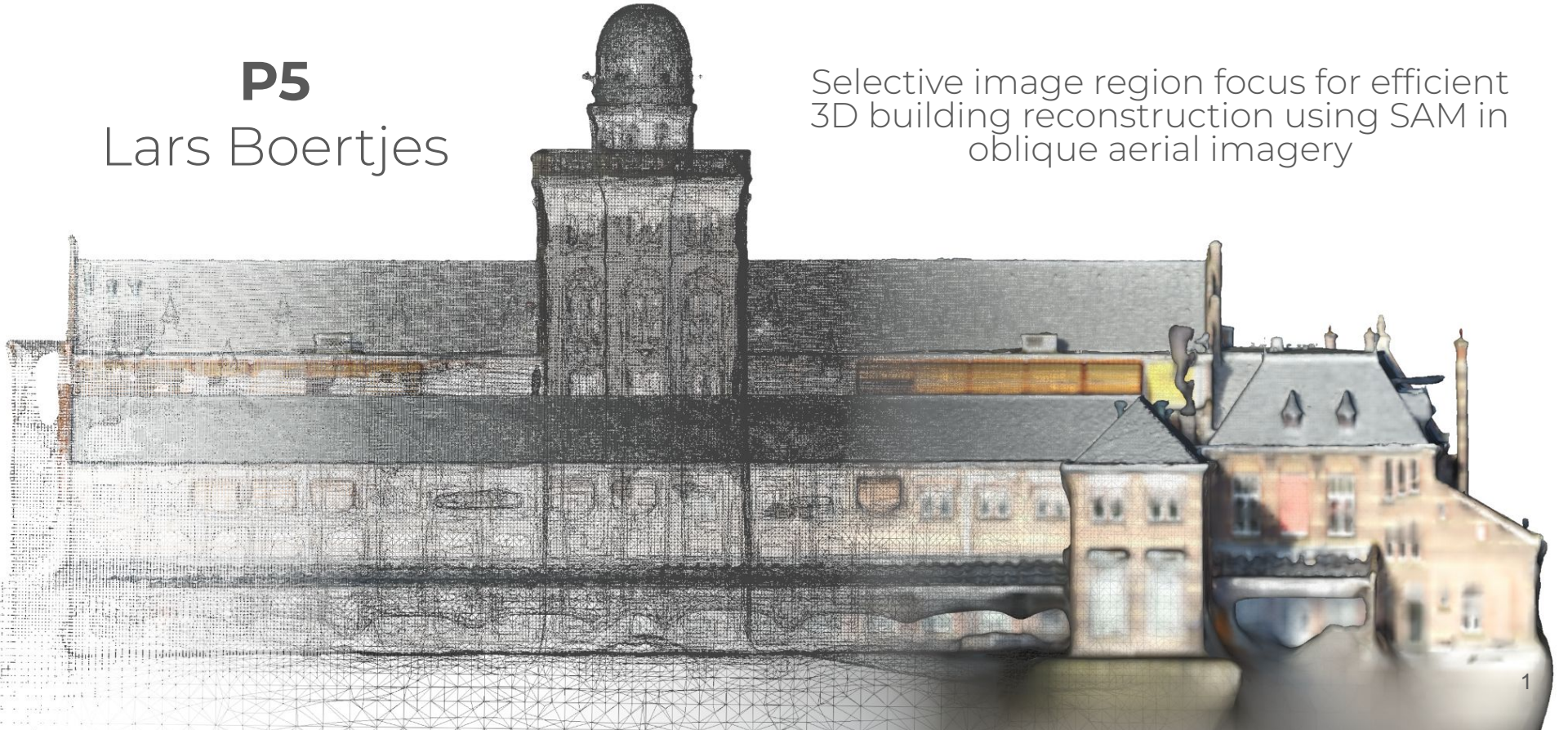
P5

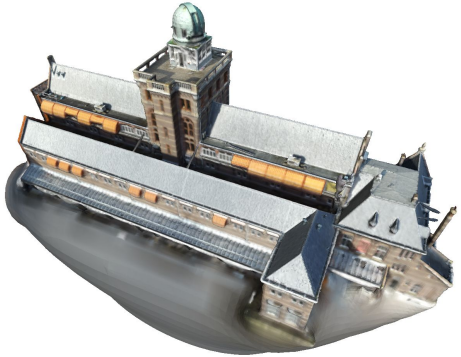
Lars Boertjes

Selective image region focus for efficient 3D building reconstruction using SAM in oblique aerial imagery



GEODELTA

Point clouds



3D meshes

Oblique aerial photos

Meta

Segment Anything Model

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?

GOAL: improve oblique -> 3D mesh (existing techniques) using SAM

Oblique imagery

3D mesh

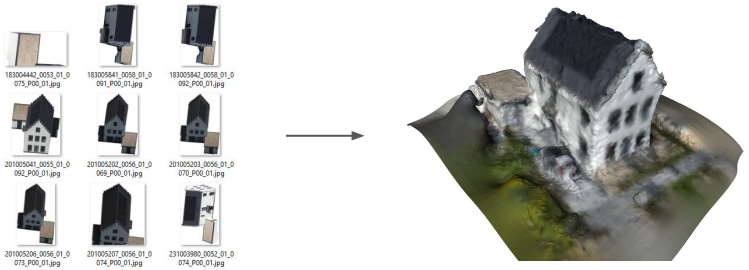
SAM

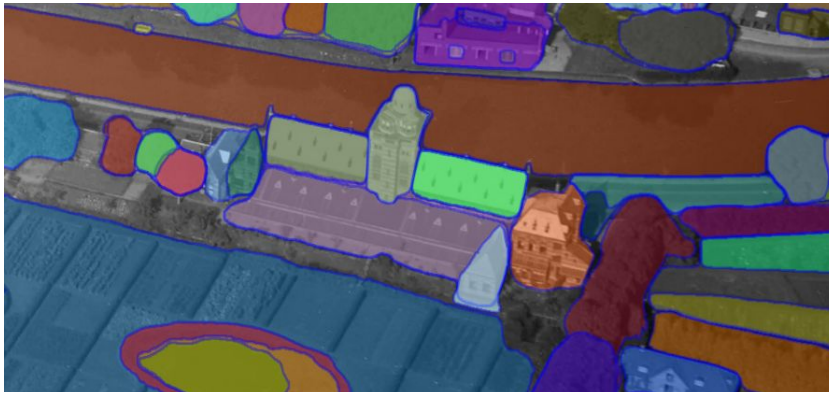
Phase I

Extracting buildings from oblique photos

Phase II

Making efficient meshes from building segments





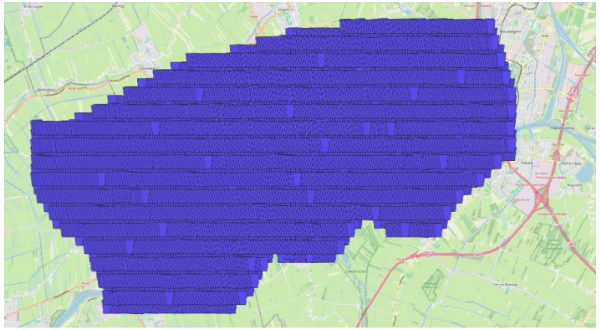
Phase I

Extracting buildings from oblique photos

Around 2,500 oblique photos

+- 40 buildings/image

BAG dataset



flight plans

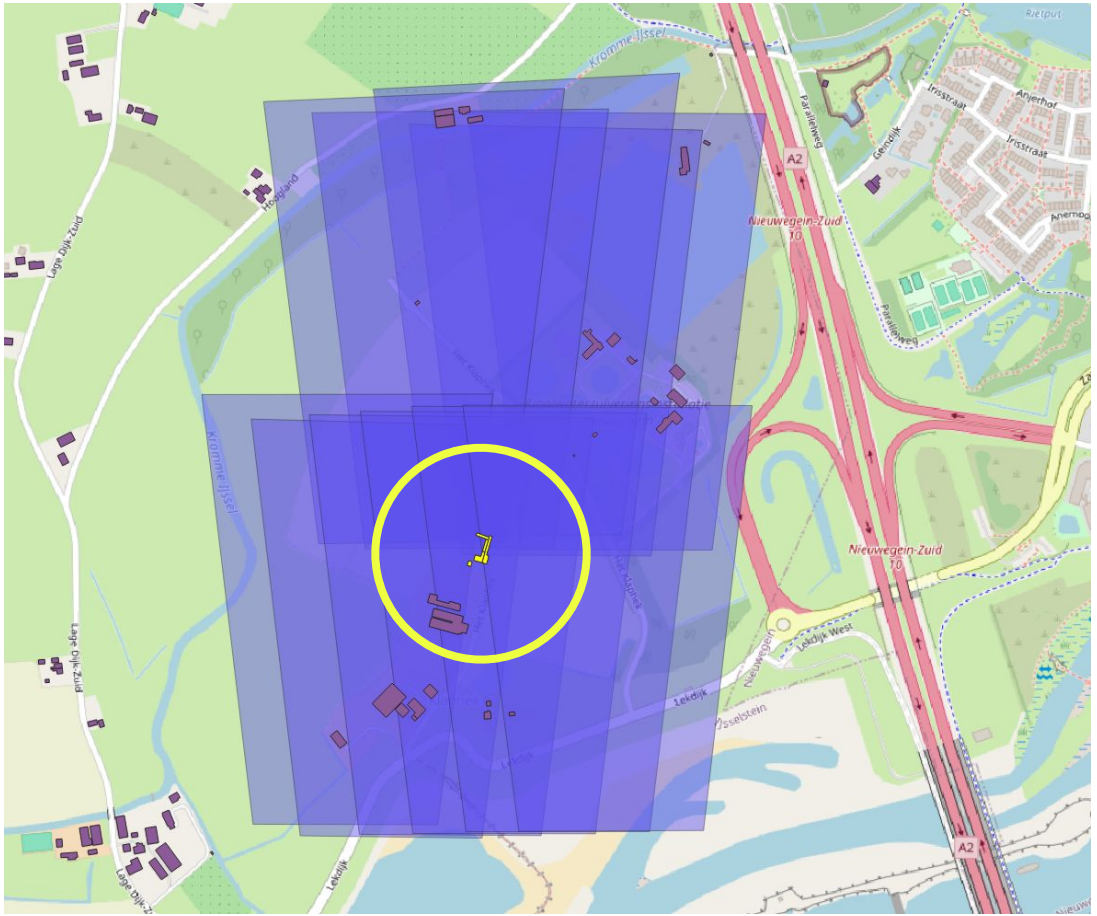
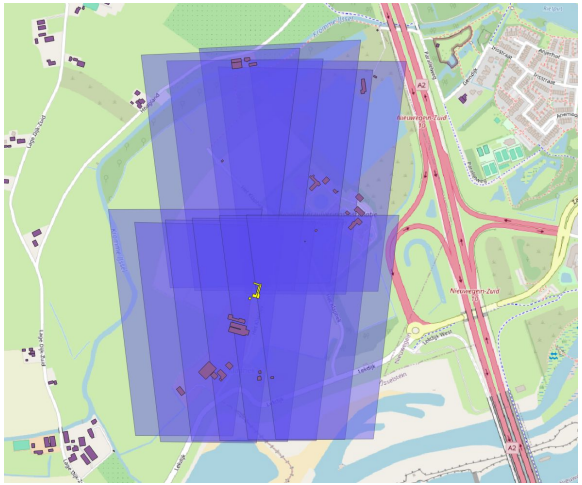
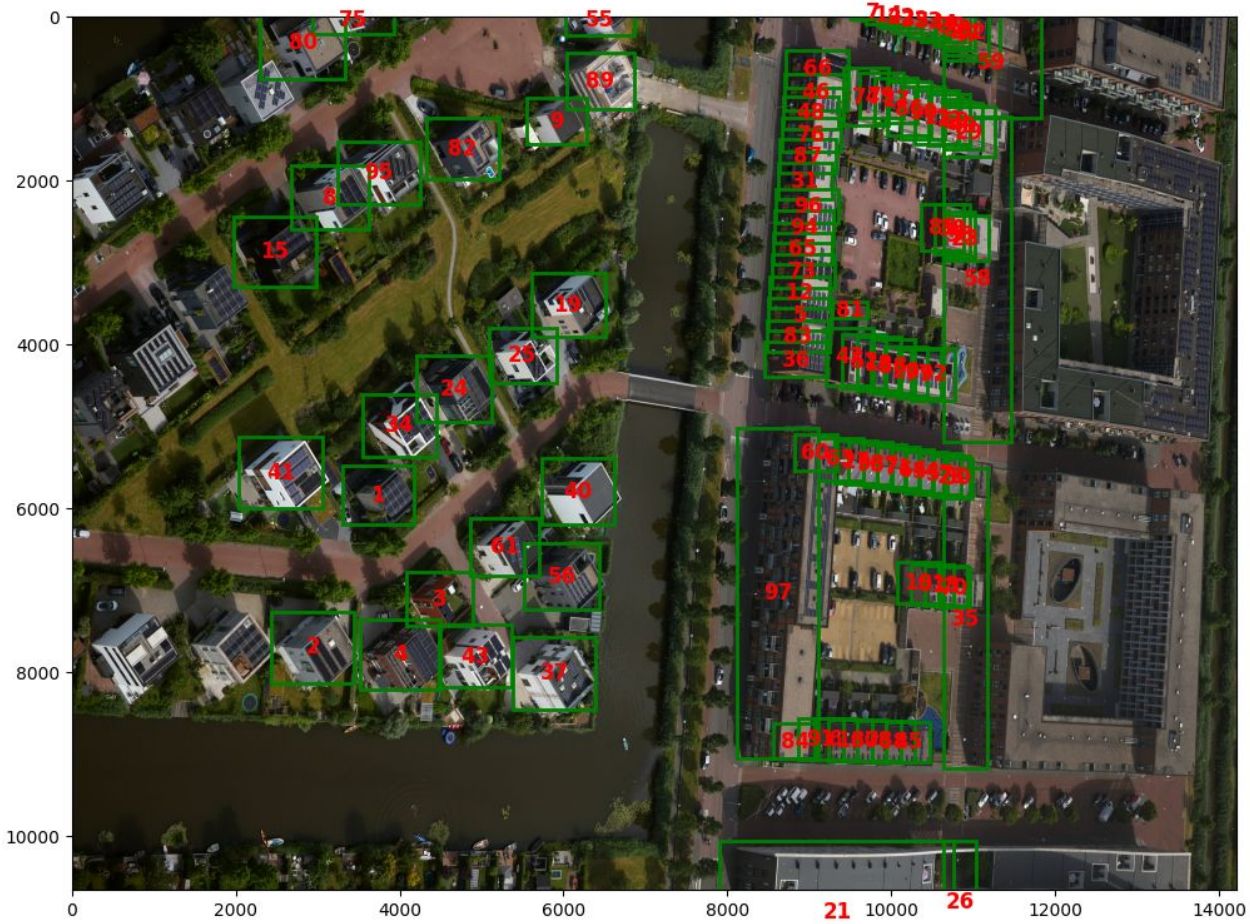


image to building correspondence



CameraId	PixelSize [µm]	Width [px]	Height [px]	Focal [mm]	PPX [mm]	PPY [mm]
Back	3.76	14204	10652	108.694	-0.045	-0.048
Fwd	3.76	14204	10652	108.752	0.06	-0.154
Left	3.76	14204	10652	108.836	-0.146	0.227
Right	3.76	14204	10652	108.807	0.033	-0.076
ImageId	X [m]	Y [m]	Z [m]	Omega [deg]	Phi [deg]	Kappa [deg]
262000203_0067_01_0110_P00_01	131187.73650	457238.56580	427.39680	-0.20540	44.80360	269.42430
262000204_0067_01_0109_P00_01	131137.40660	457238.13440	429.09200	-0.19240	44.81710	269.36400
262000205_0067_01_0108_P00_01	131087.09660	457237.63790	431.21260	-0.22510	44.83760	269.44480
262000206_0067_01_0107_P00_01	131036.71030	457237.31560	433.30000	-0.13140	44.85400	269.47040
262000207_0067_01_0106_P00_01	130986.29360	457237.65140	435.29080	-0.56930	44.80520	269.04450

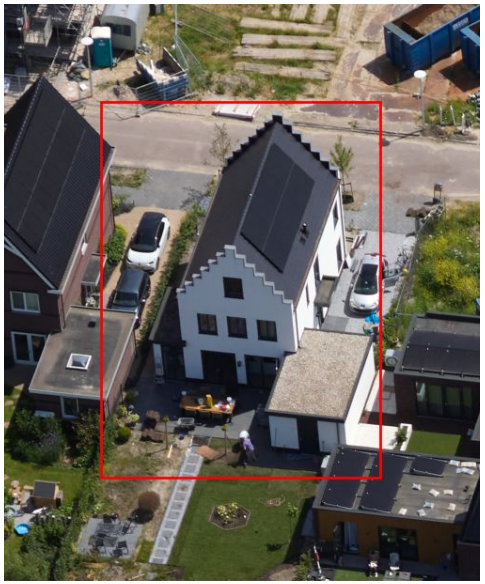
Obliques orientation data


0
2000
4000
6000
8000
10000
0
2000
4000
6000
8000
10000
12000
14000
21
26

Phase I

Using the reprojected bounding box as input prompt for SAM

Obliques



2D image prompt

Segmented mask

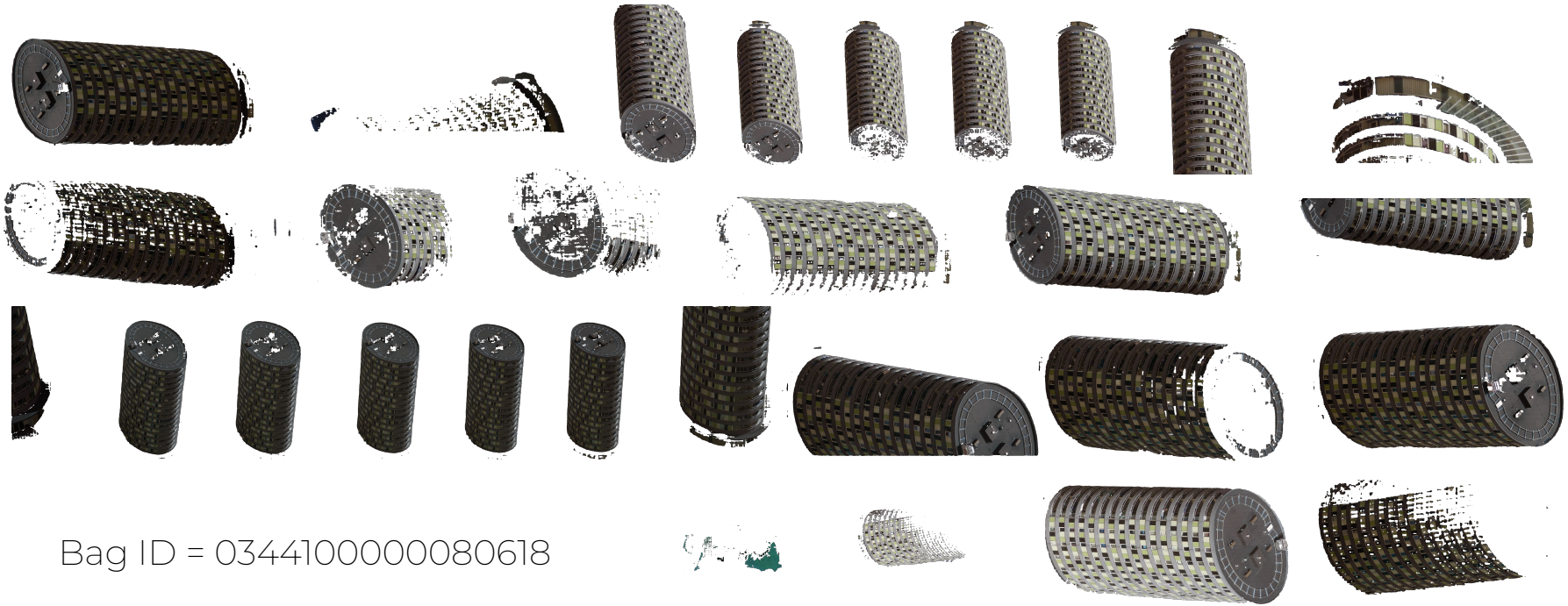
Phase I

All the mask belonging to one building

Bag ID = 0344100000157440

Phase I

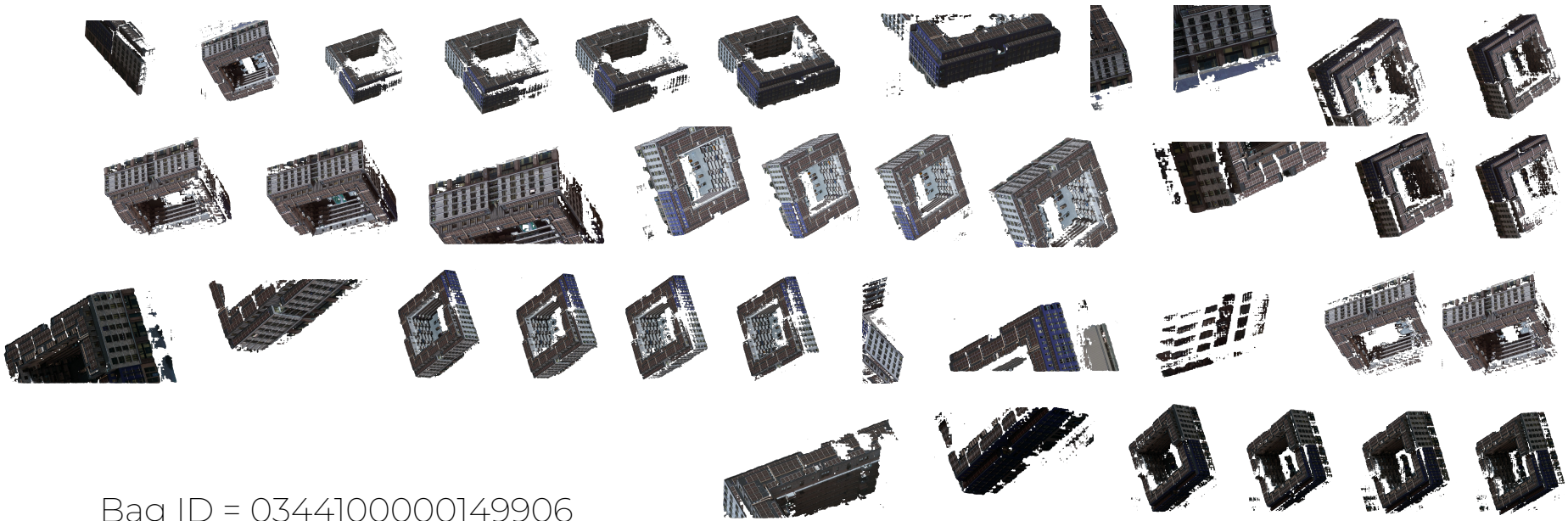
All the mask belonging to one building



Bag ID = 0344100000080618

Phase I

All the mask belonging to one building



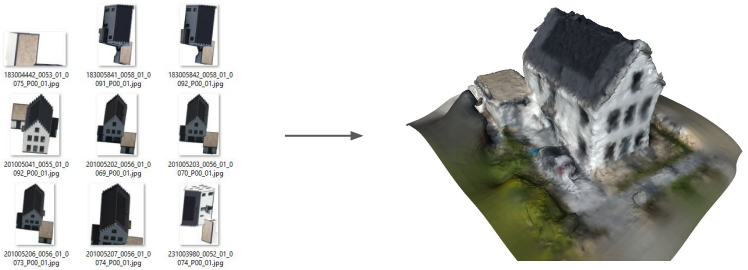
Bag ID = 0344100000149906

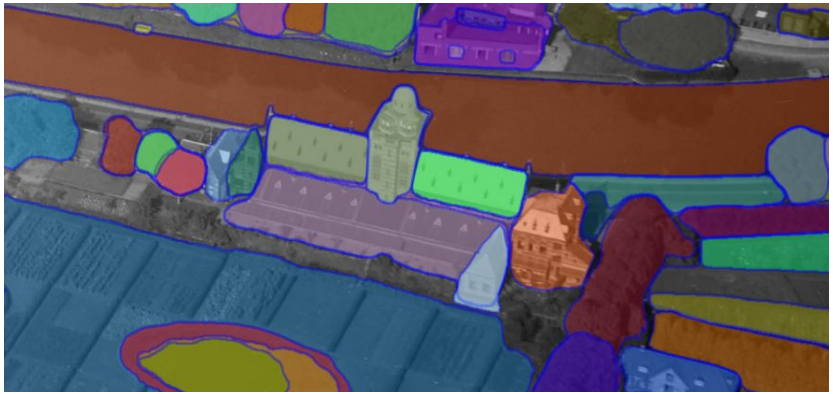
Phase I

Extracting buildings from oblique photos

Phase II

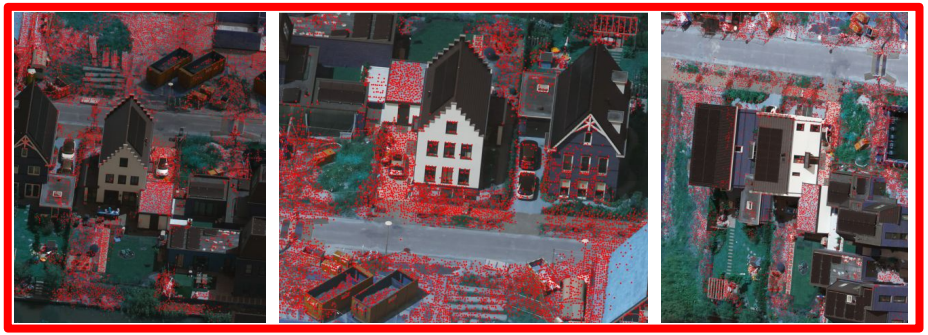
Making efficient meshes from building segments



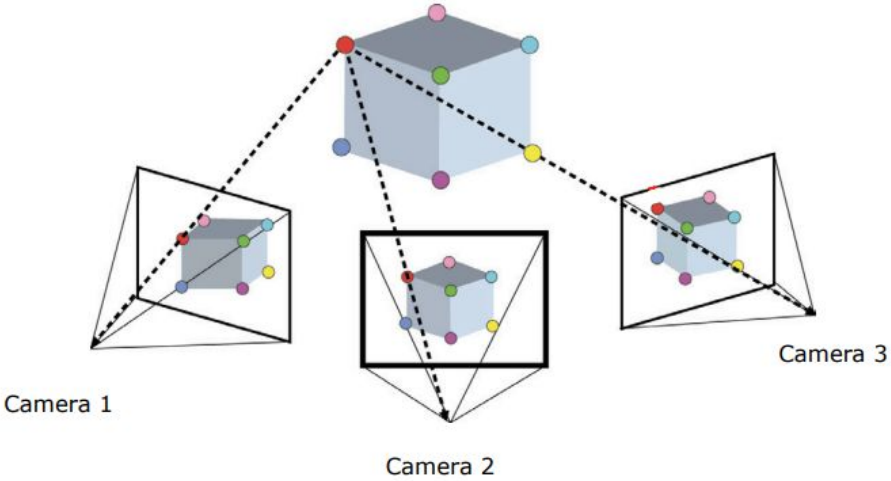


1. Input photos

2. Feature extraction & matching



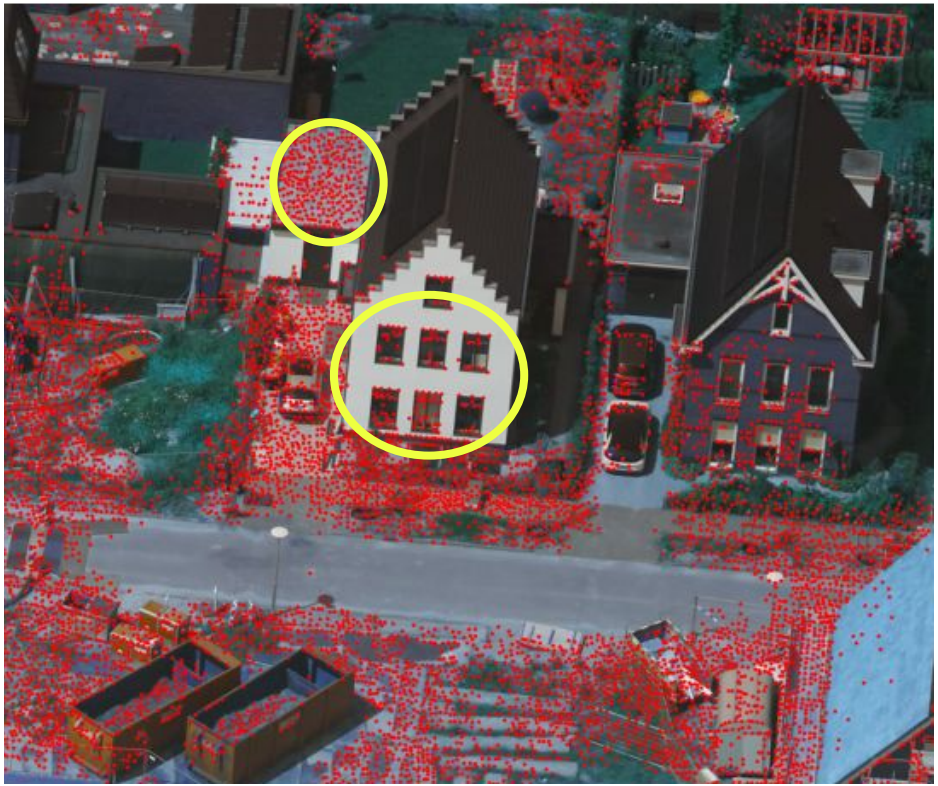
3. Sparse 3D reconstruction



4. Dense 3D reconstruction

5. Mesh creation

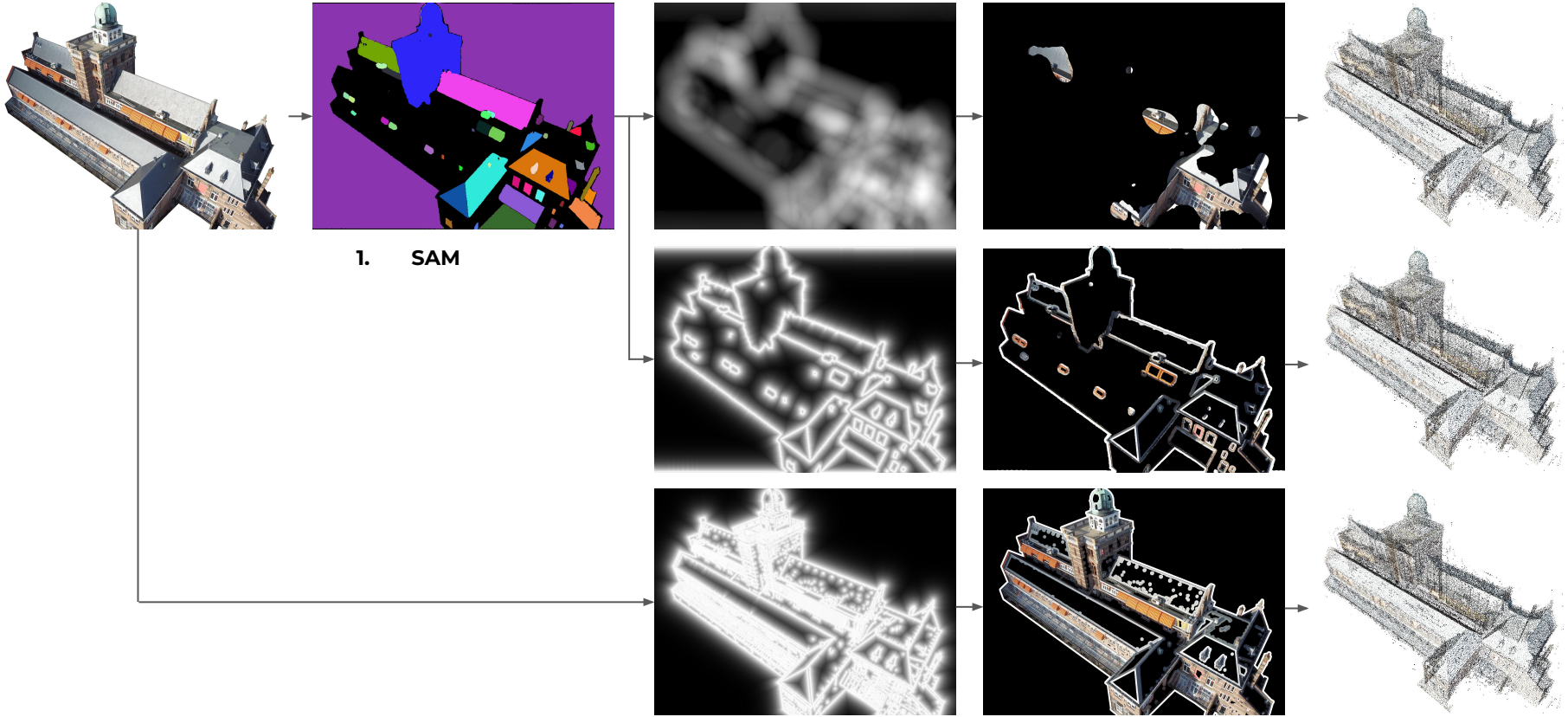
PROBLEM: key points are extracted without connection to mesh quality



GOAL: Guide algorithm on where to spend resources

Phase II

Selecting important regions in the images for reconstruction



1. SAM

2. Importance maps

3. Thresholding

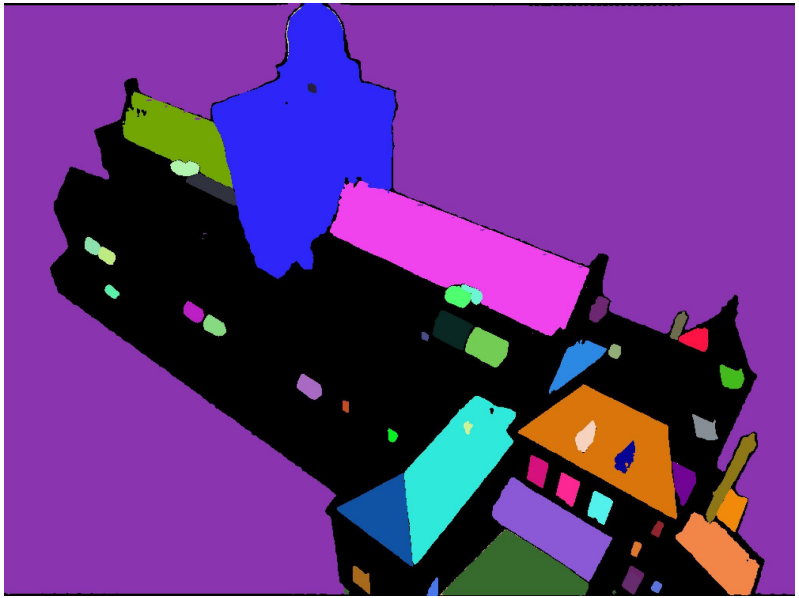
4. Reconstruction

Phase II

SAM kernel entropy

Importance map

SAM segmentation



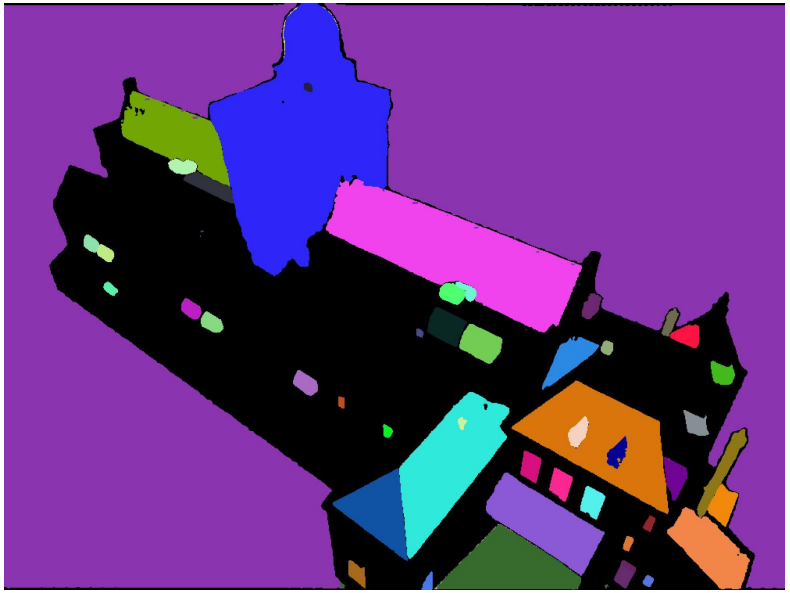
Importance is where there is most uncertainty about mask distribution

Phase II

SAM edge distance

Importance map

SAM segmentation

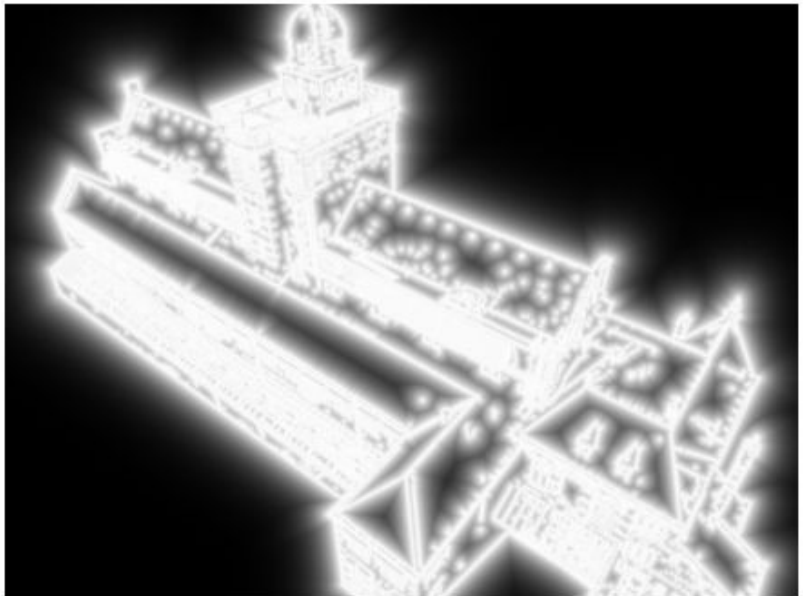


Importance is near the edges of SAM masks

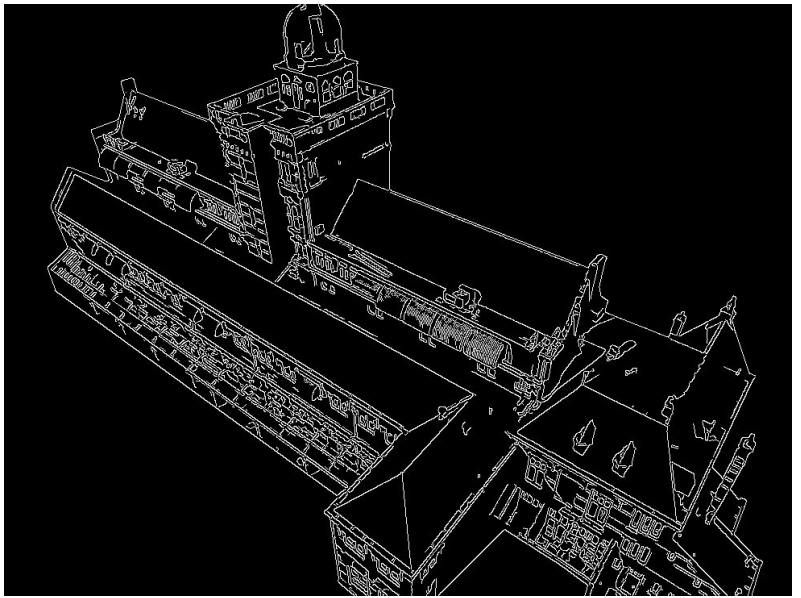
Phase II

Canny edge detection

Importance map



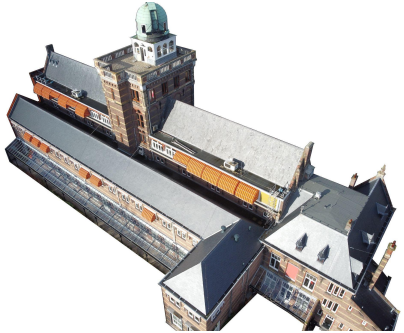
Canny edge detection



Importance is near the visual edges

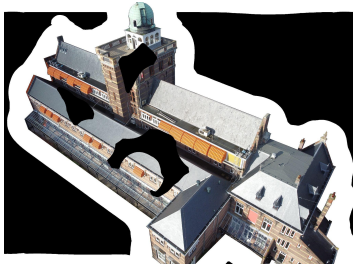
Phase II

thresholding

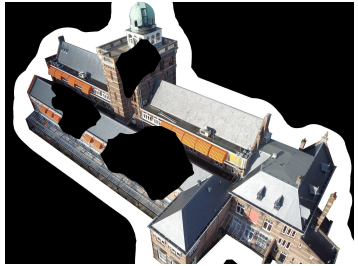


Input image

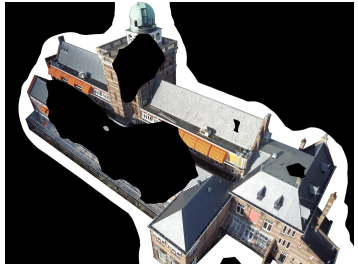
Importance map [0-1.0]
SAM kernel entropy



0.1



0.2



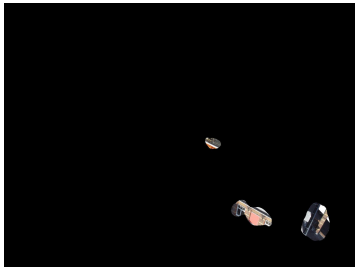
0.3

0.4

0.5

0.6

0.7

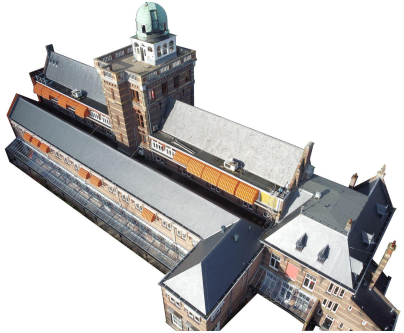


0.8

0.9

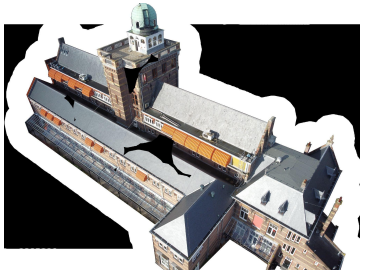
Phase II

thresholding

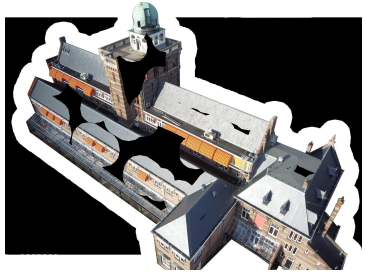


Input image

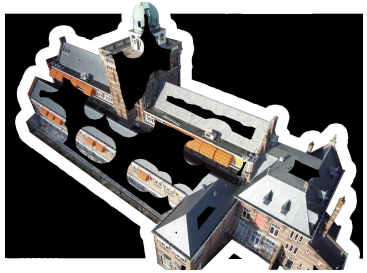
Importance map [0-1.0]
SAM edge distance



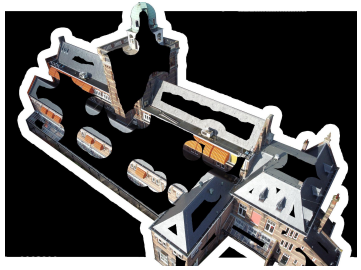
0.1



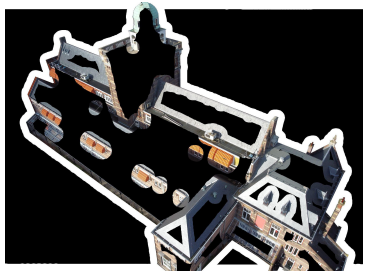
0.2



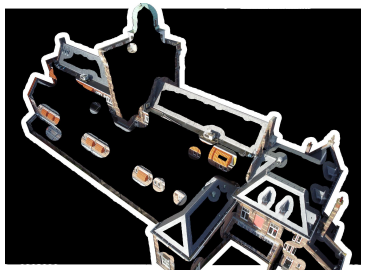
0.3



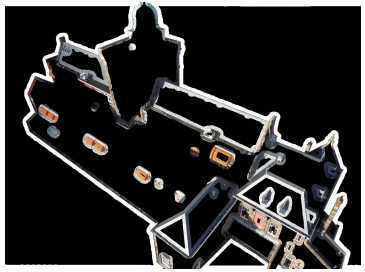
0.4



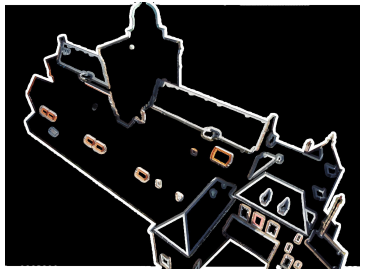
0.5



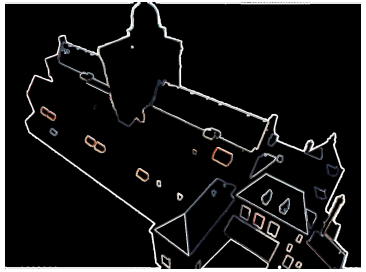
0.6



0.7



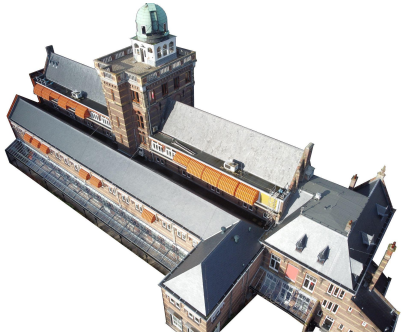
0.8



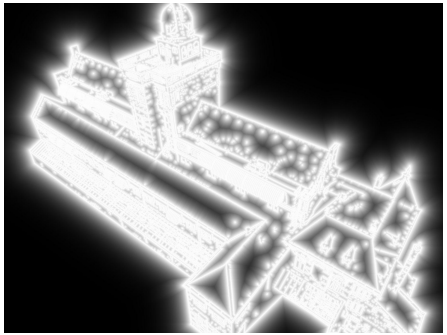
0.9

Phase II

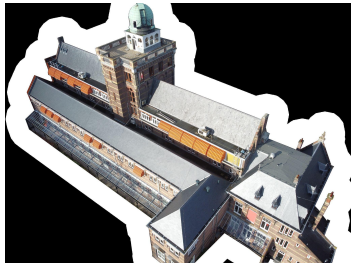
thresholding



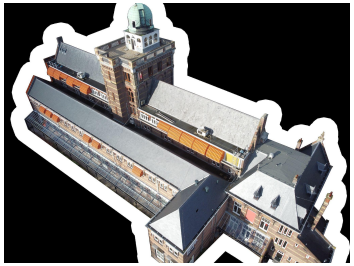
Input image



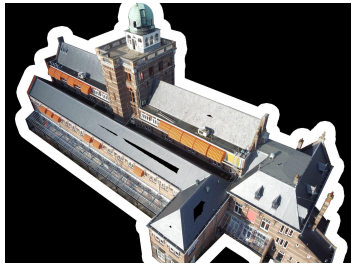
Importance map [0-1.0]
SAM kernel entropy



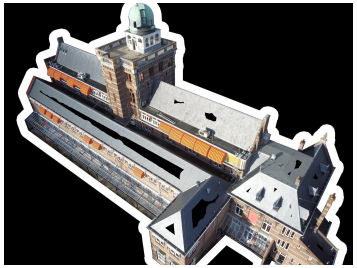
0.1



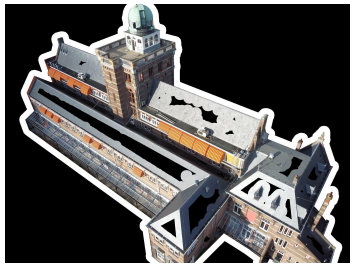
0.2



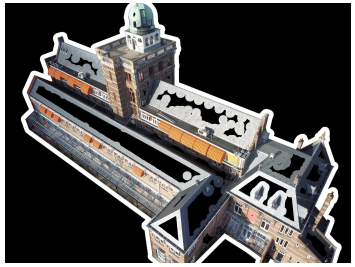
0.3



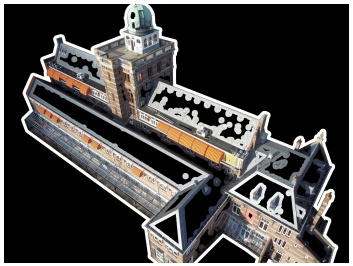
0.4



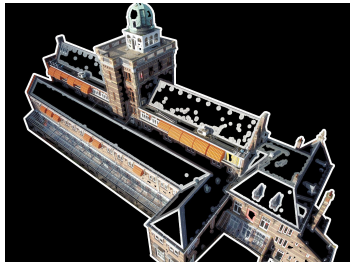
0.5



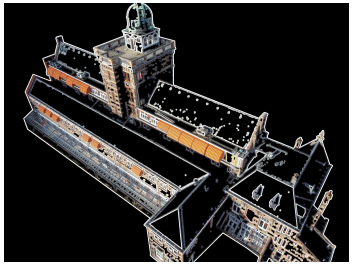
0.6



0.7



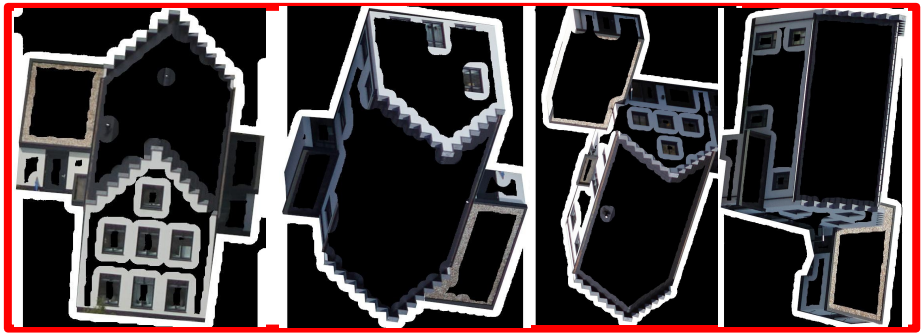
0.8



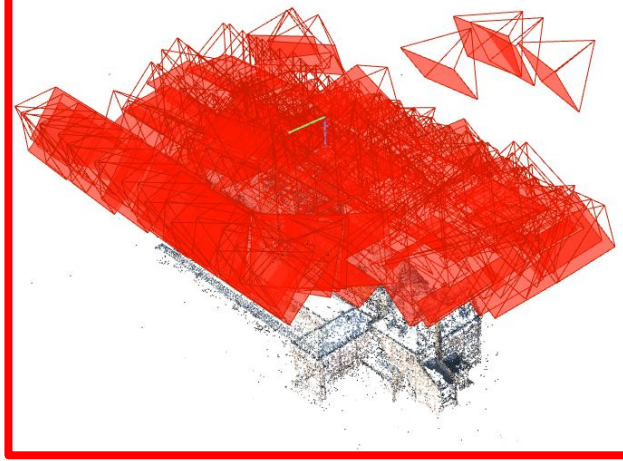
0.9

1. Input photos

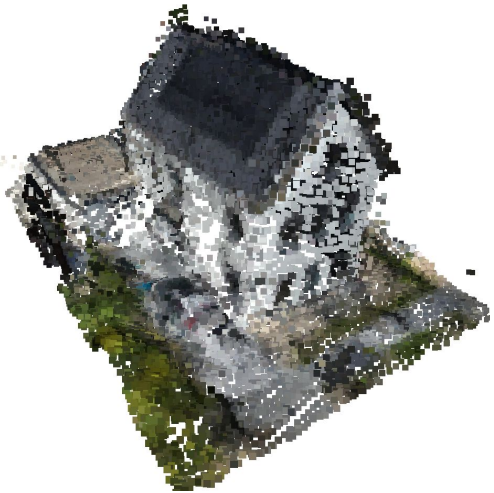
2. Feature extraction & matching



3. Sparse 3D reconstruction



4. Dense 3D reconstruction



5. Mesh creation

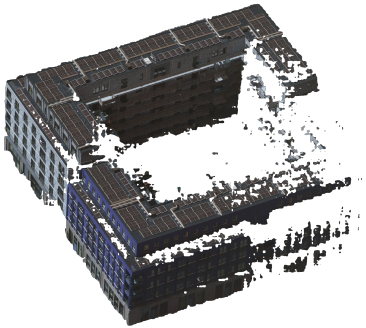
Phase II

Reconstruction

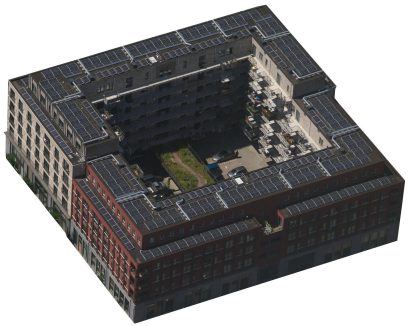
SAM

Manual

SAM



SAM

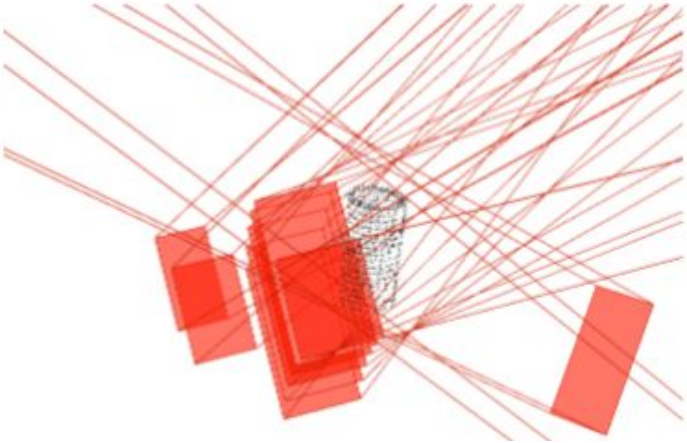


Manual

5 different datasets
3 different importance maps
9 different thresholds
135 configurations for reconstruction

Phase II

Reconstruction

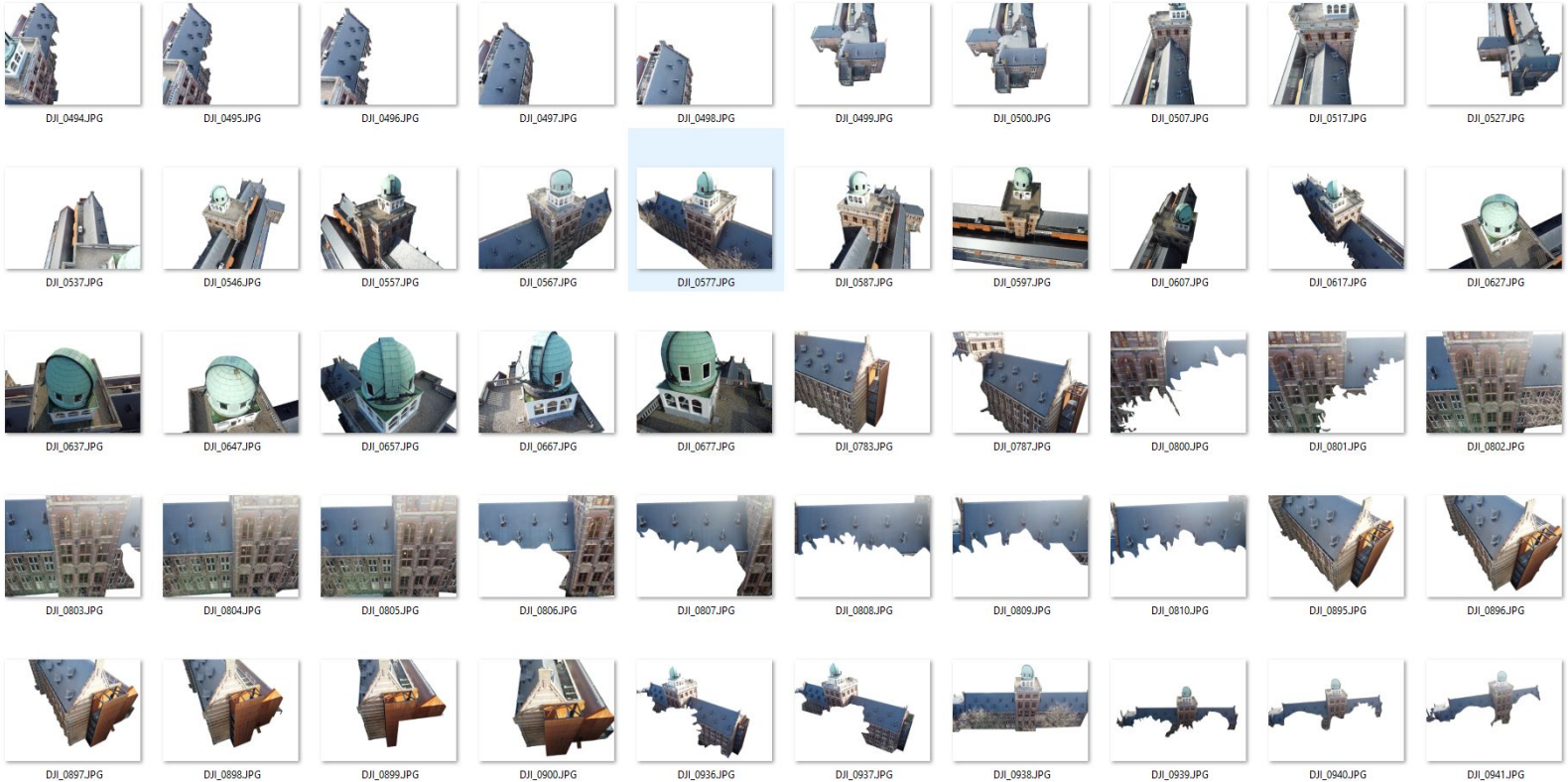


Only one reconstruction success

Not enough context and overlap for good reconstruction

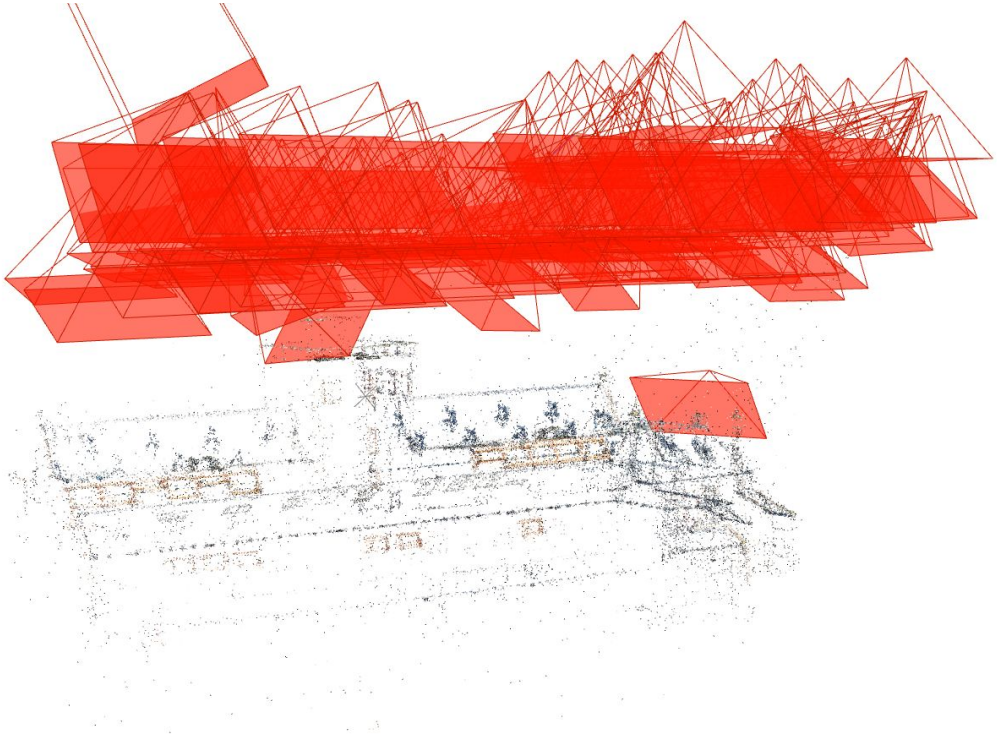
Phase II

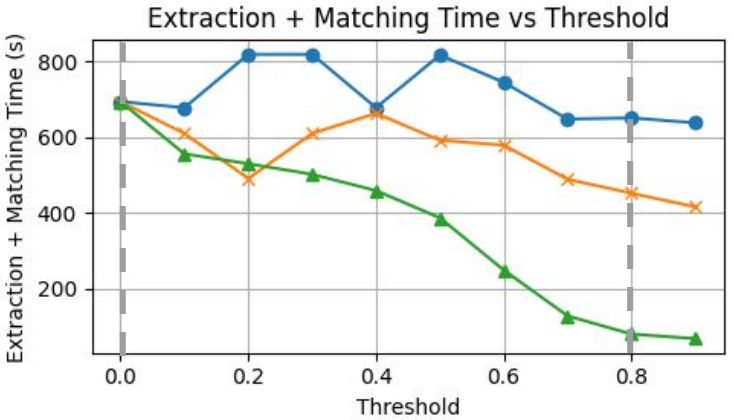
Geodelta office drone dataset

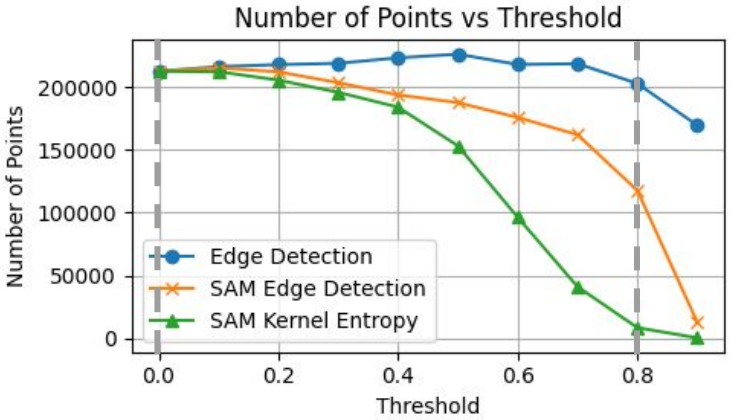


Phase II

Geodelta office drone dataset

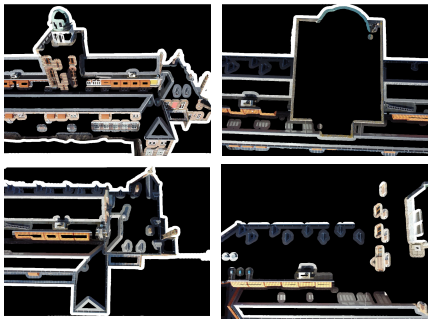




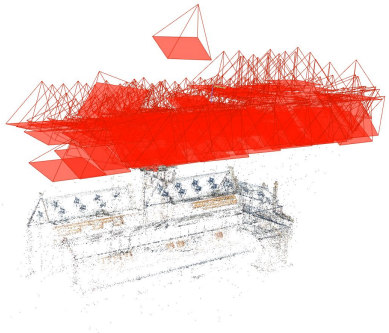


Phase II

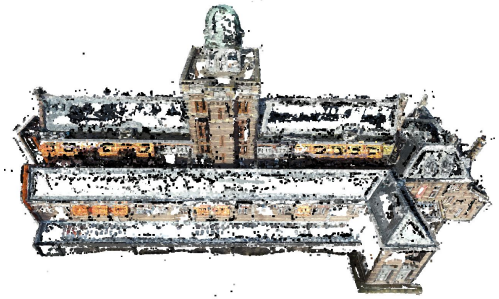
Evaluating mesh quality



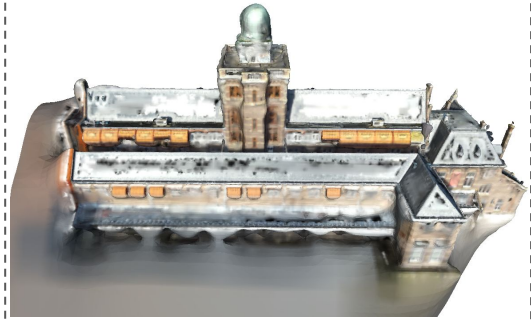
SAM edge distance 0.6



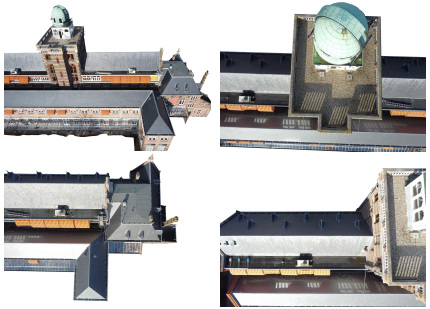
sparse reconstruction



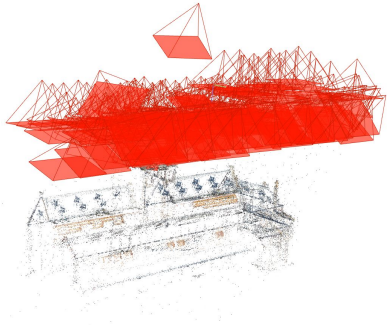
dense reconstruction



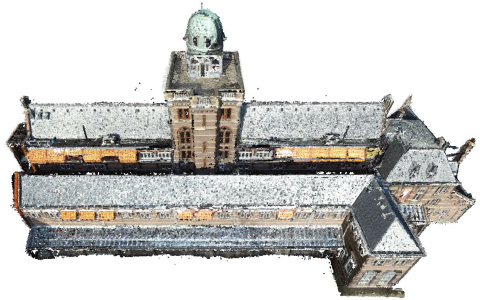
mesh



Full images



sparse reconstruction

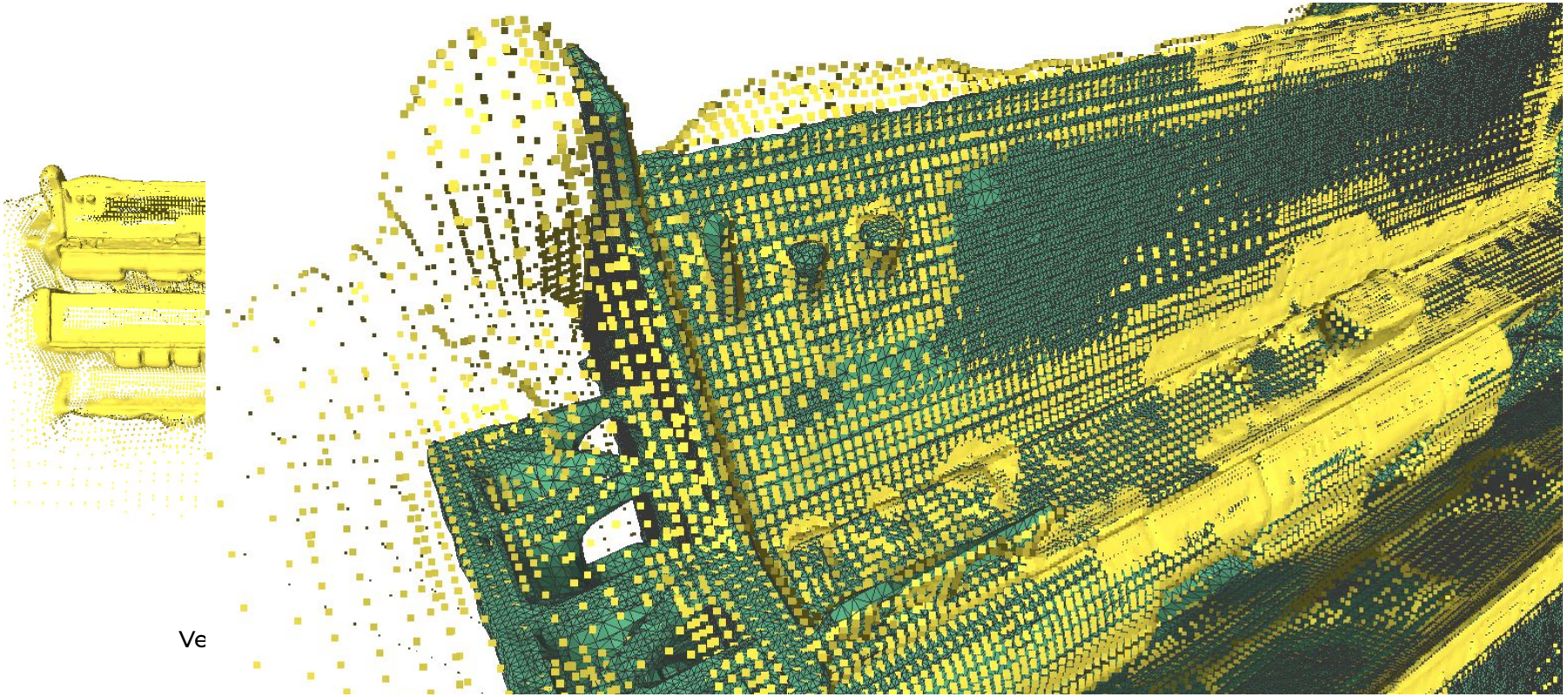


dense reconstruction

Ground truth

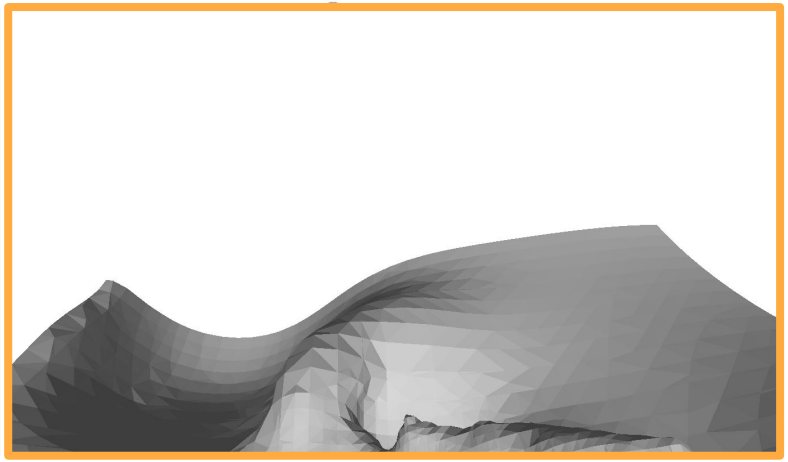
Phase II

Evaluating mesh quality

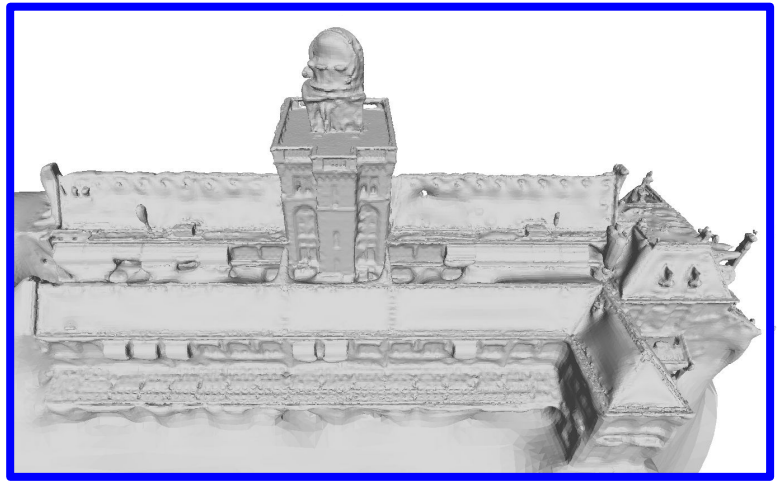


Ve

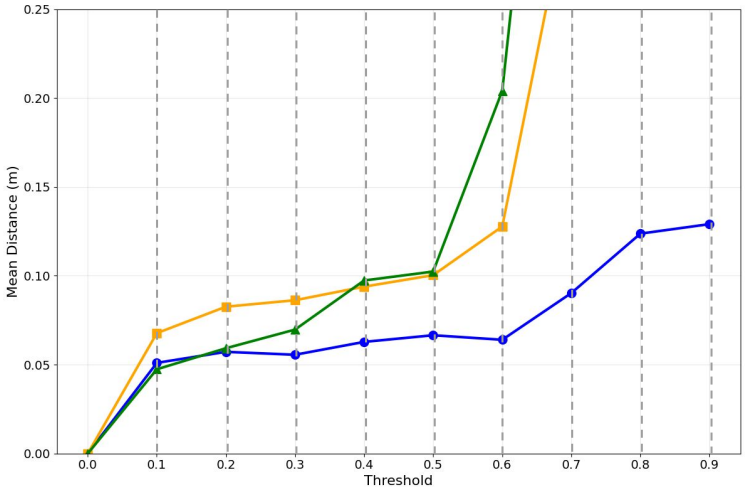
SAM kernel entropy



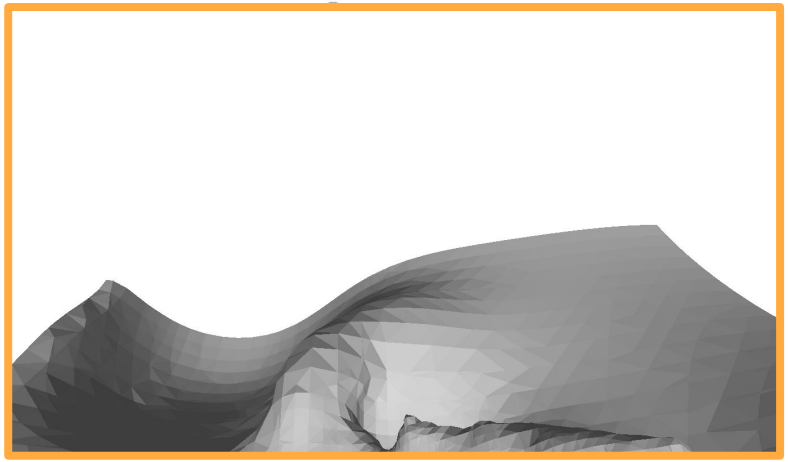
SAM edge distance



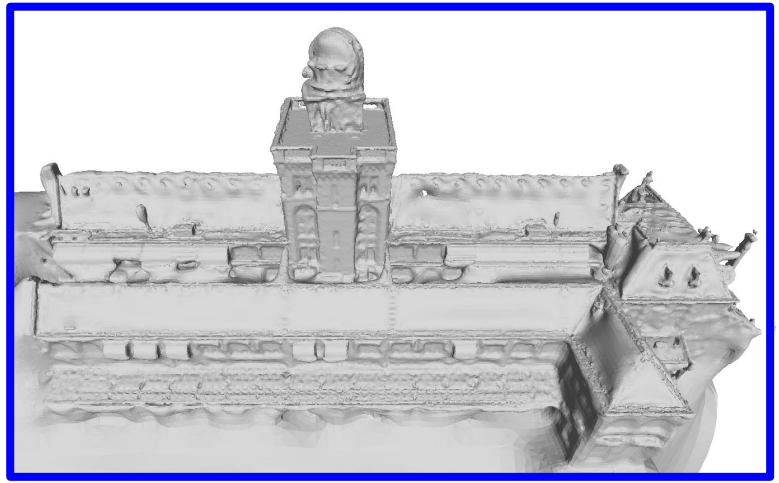
Canny edge distance



SAM kernel entropy



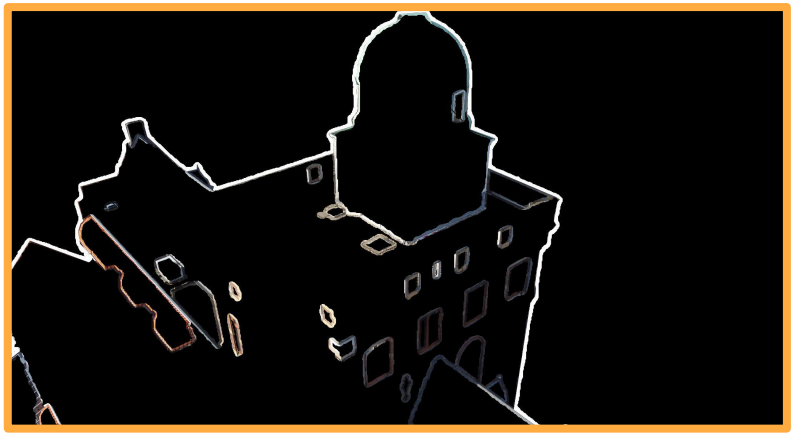
SAM edge distance

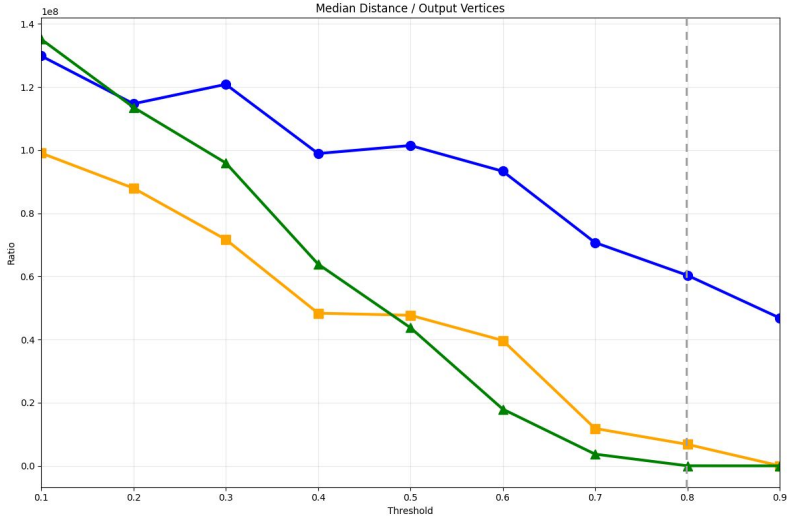


Canny edge distance

SAM Kernel Entropy		SAM Edge Distance Entropy		Canny Edge Entropy	
Points	% Savings	Points	% Savings	Points	% Savings
7,564,981	0%	7,564,981	0%	7,564,981	0%
6,842,222	9.6%	7,339,276	3.0%	7,425,792	1.8%
6,353,980	16.0%	6,673,821	11.8%	7,248,932	4.2%
5,663,213	25.1%	5,539,212	26.8%	7,070,308	6.5%
3,785,706	50.0%	4,520,885	40.2%	6,743,709	10.9%
2,533,572	66.5%	3,522,776	53.4%	6,306,428	16.6%
1,157,186	84.7%	2,498,264	66.9%	5,996,500	20.7%
442,145	94.2%	1,728,857	77.1%	5,041,322	33.4%
53,149	99.3%	982,411	87.0%	4,484,251	40.7%
1,404	99.98%	108,029	98.6%	3,326,978	56.0%

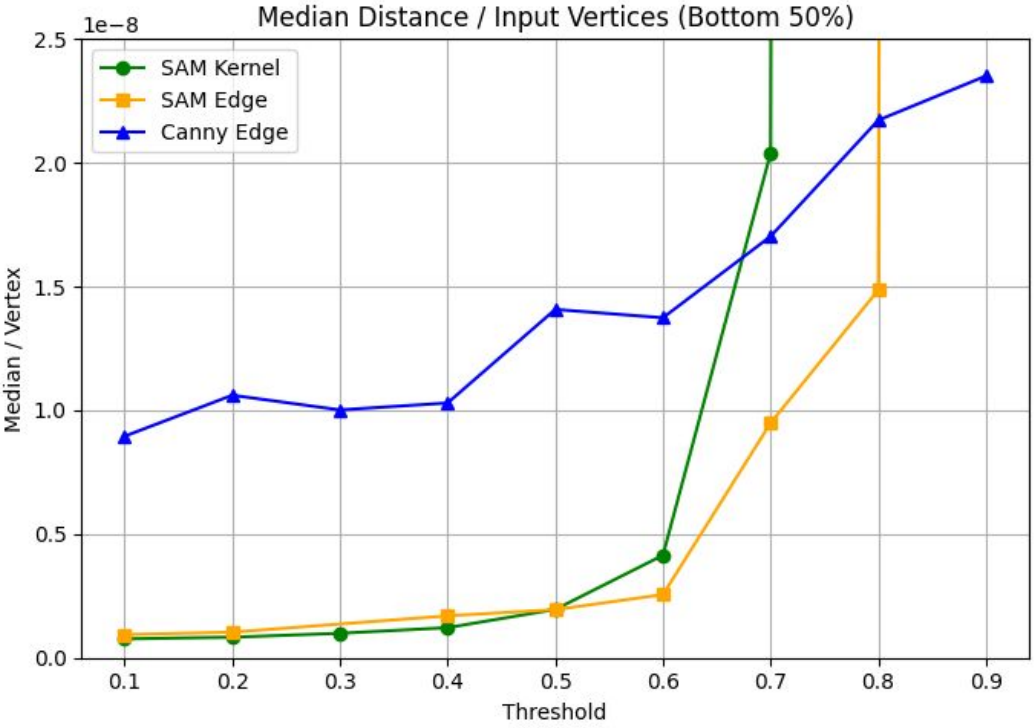
Mesh quality - output size ratio

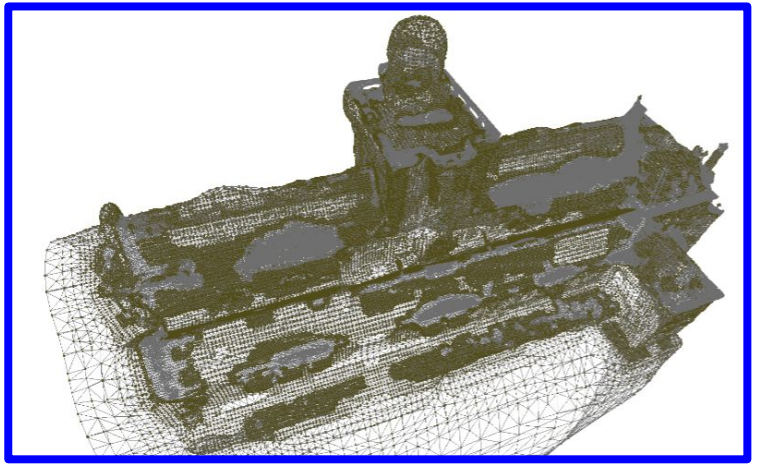




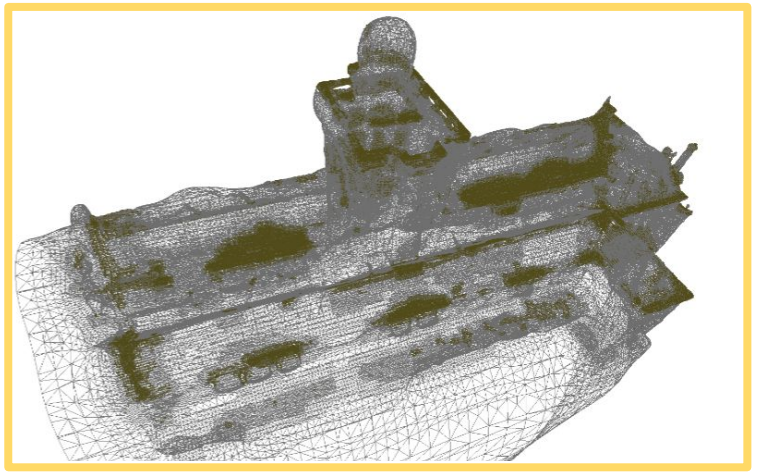
Phase II

Dense vs. sparse quality





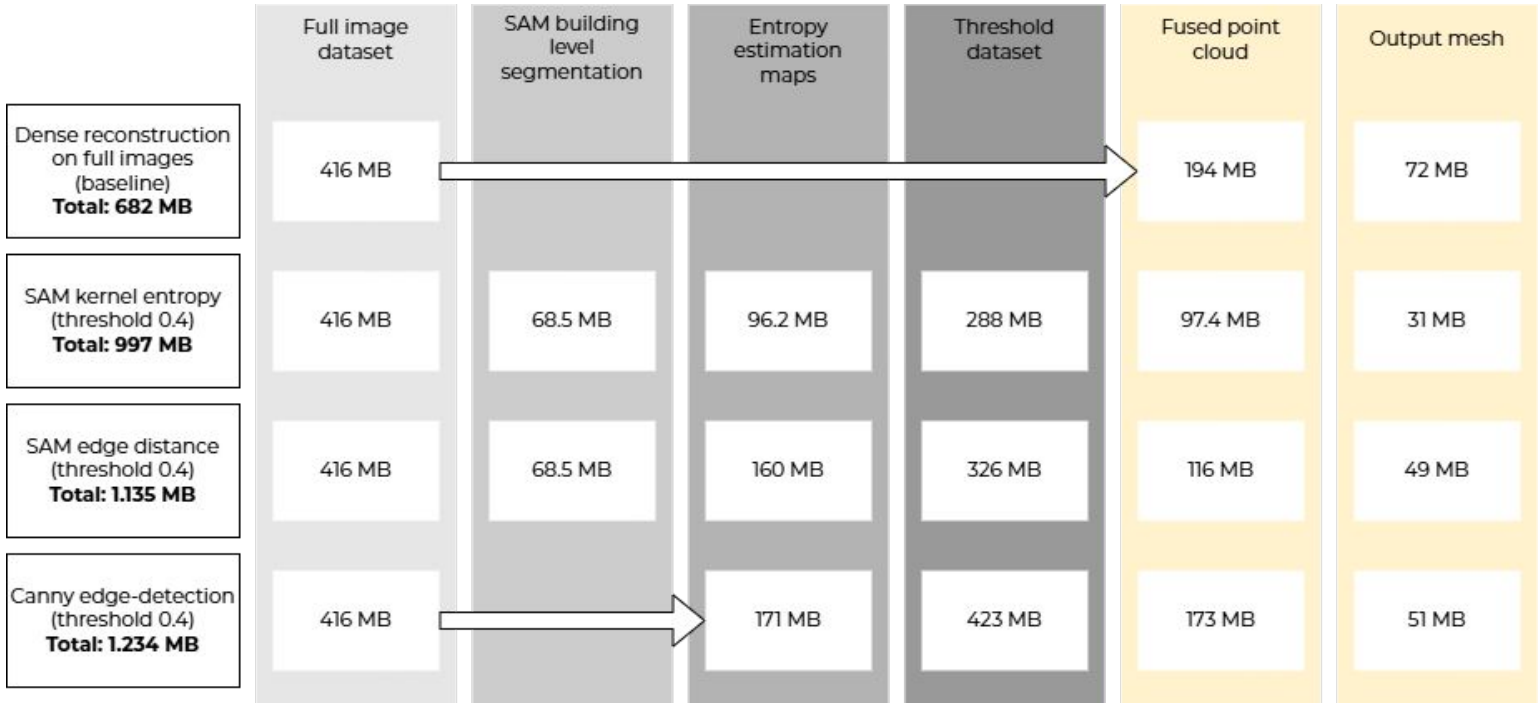
dense areas (top 50% density)

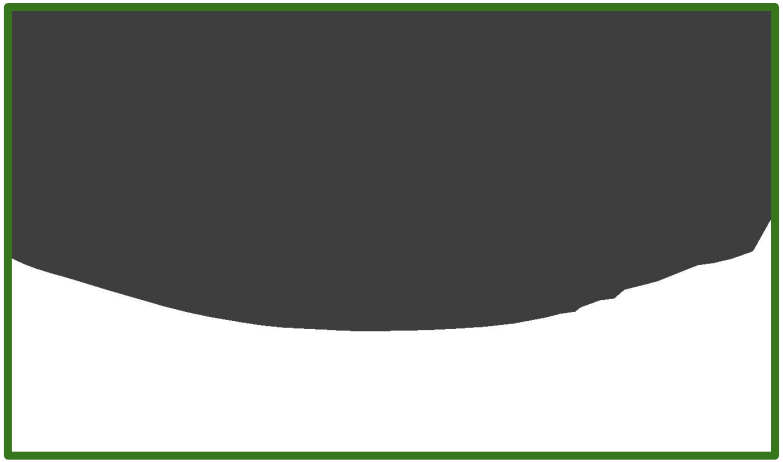


sparse areas (bottom 50% density)

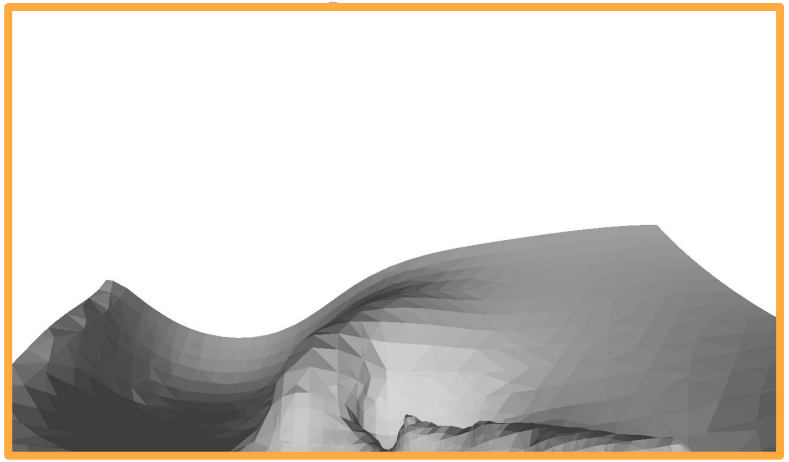
Phase II

Full memory pipeline context

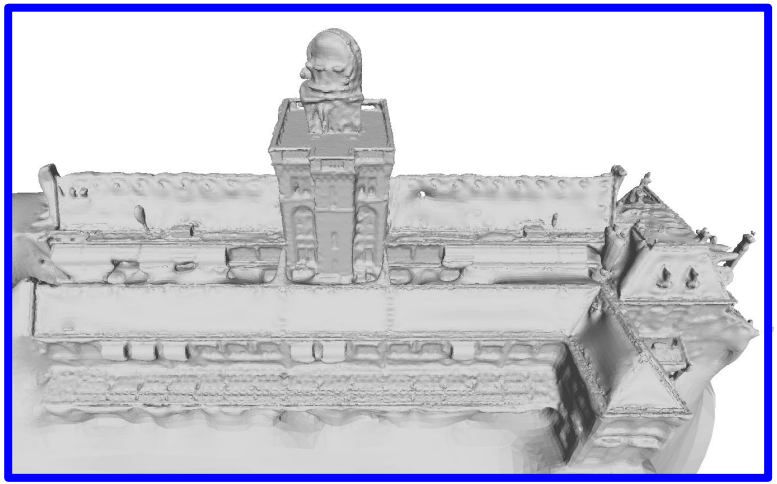




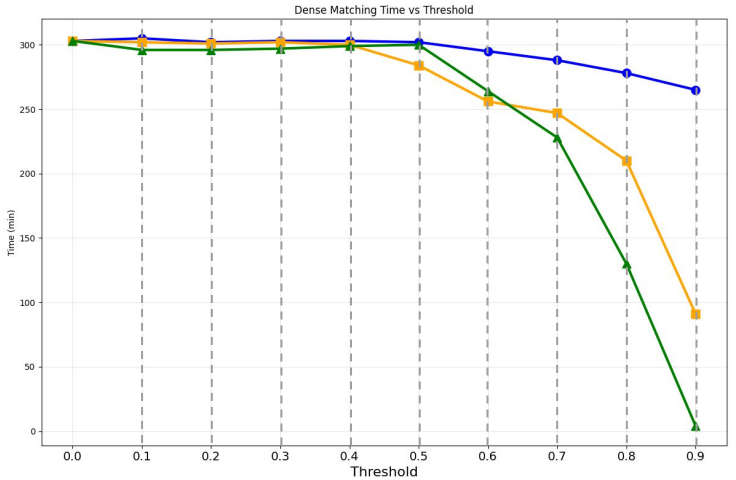
SAM kernel entropy



SAM edge distance



Canny edge distance



10. Results - mesh quality vs. runtime

Only slight reduction in runtime on dense matching, full pipeline increases

	SAM building level segmentation	Entropy estimation maps generation	Thresholding of image dataset	Dense matching
Dense reconstruction on full images (baseline) **Total: 5h3m**				5h3m
SAM kernel entropy (threshold 0.4) **Total: 12h24m44s**	6h32m14s	51m20s	2m10s	4h59m
SAM edge distance (threshold 0.4) **Total: 12h8m43s**	6h32m14s	34m19s	2m10s	5h0m
Canny edge-detection (threshold 0.4) **Total: 5h7m55s**		2m45s	2m10s	5h3m

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?

Oblique imagery

3D mesh

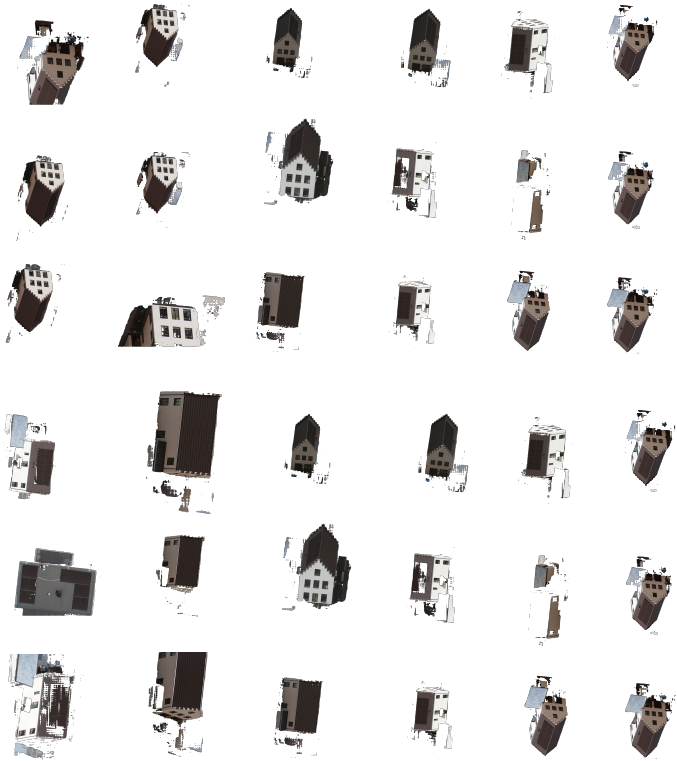
SAM

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?

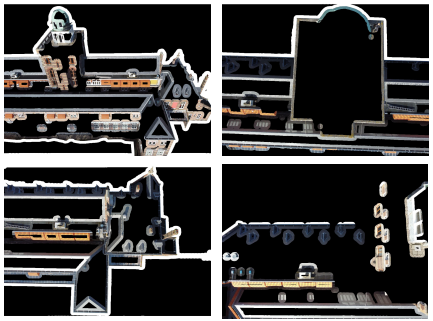


Around 250,000 oblique photos

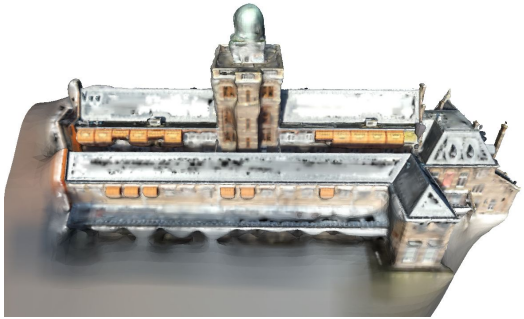
+400 buildings/image



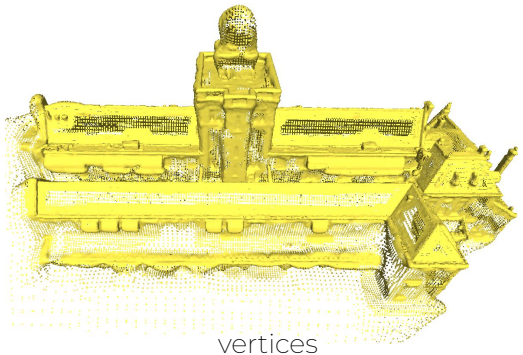
How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?



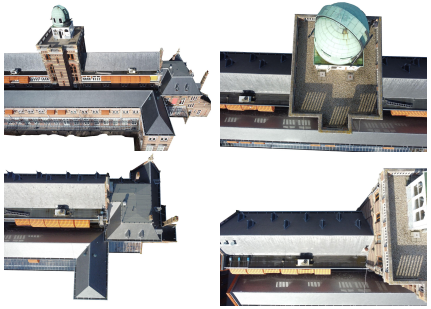
SAM edge distance 0.6



mesh

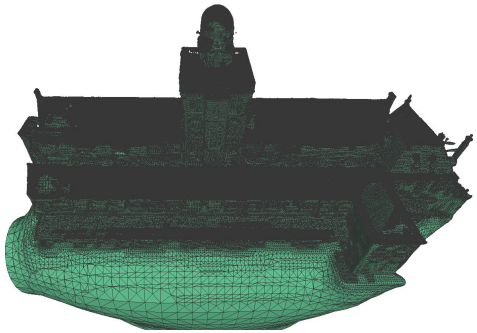


vertices



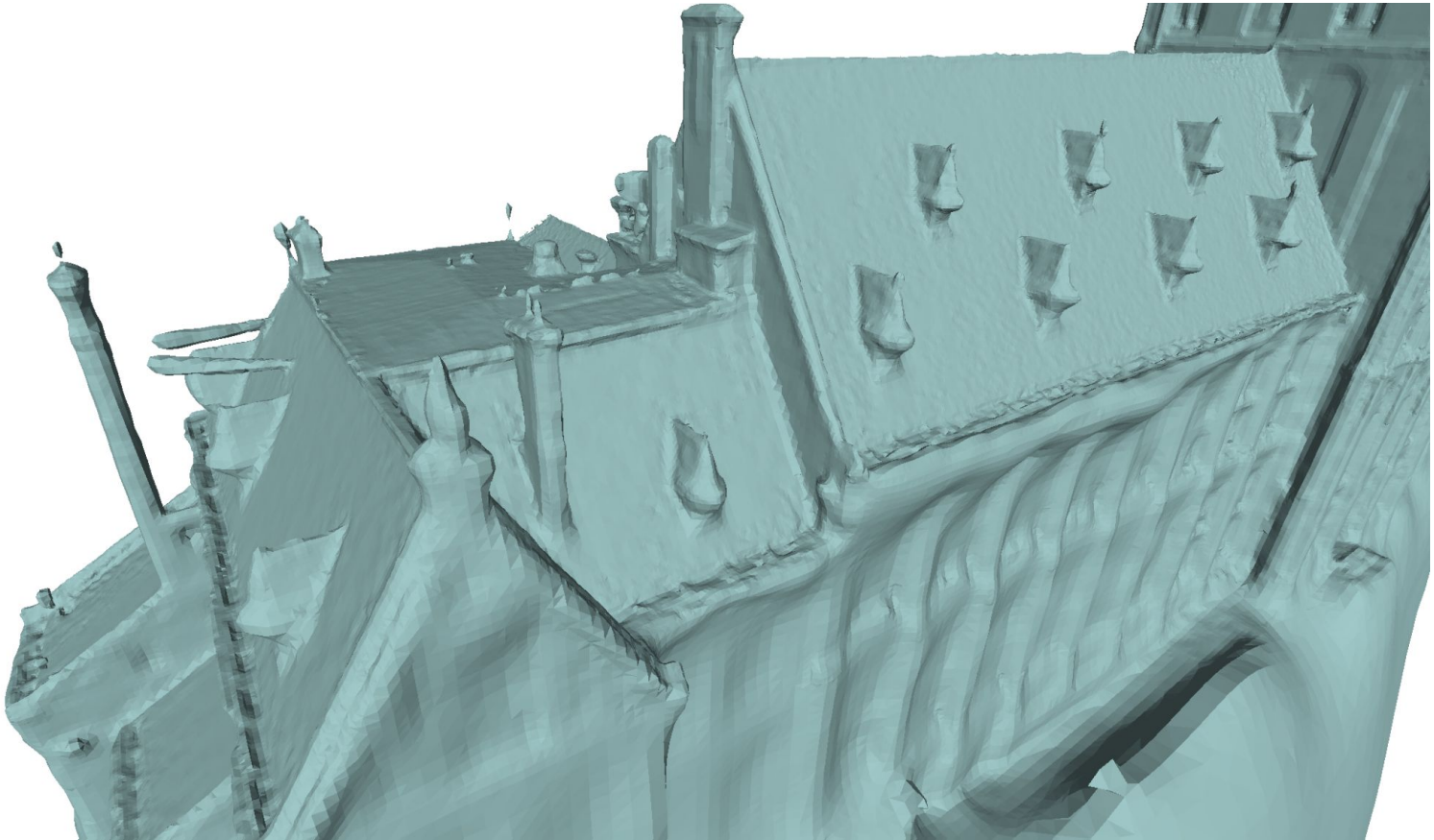
Full images

Ground truth



faces

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?



Threshold = 0.9

Threshold = 0.8

Threshold = 0.7

Threshold = 0.6

Threshold = 0.5

Threshold = 0.4

Threshold = 0.3

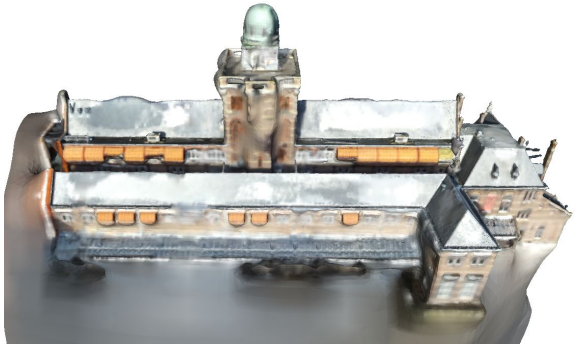
Threshold = 0.2

Threshold = 0.1

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?

canny 0.5

sam edge 0.5



sam kernel 0.5

How can *SAM*-based segmentation and importance estimation in *oblique aerial imagery* be used to improve the efficiency of *3D mesh reconstruction*?

P5

Lars Boertjes

Thanks for listening!