Reconstructing the World’s Museums

Jianxiong Xiao

MIT

Yasutaka Furukawa

Google

Ciao, Firenze!

Hello, Florence!

Maps

Photorealistic Maps

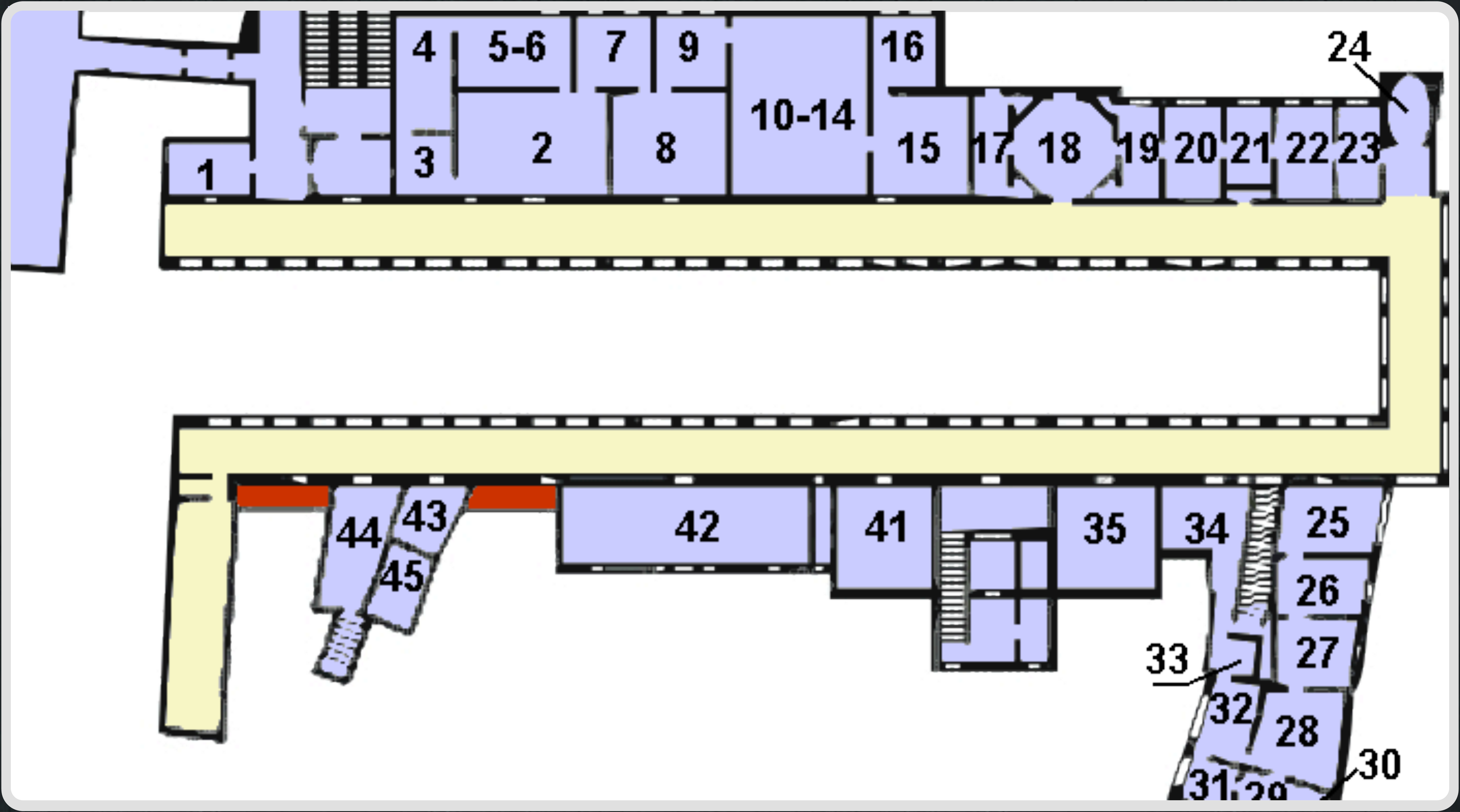
That is great!

What about indoors?

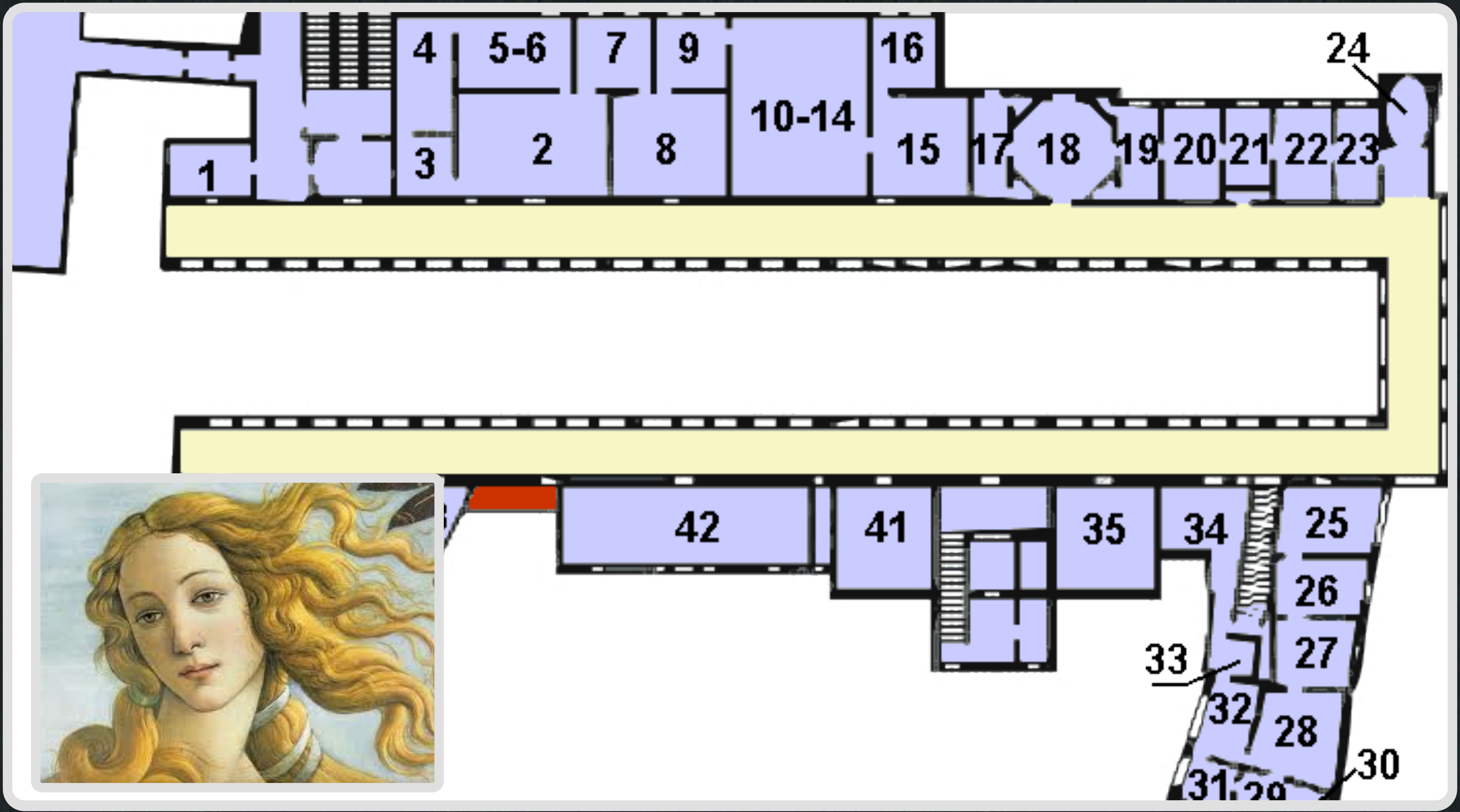
Florence = Renaissance Center of Art

Uffizi museum = one of the most famous art galleries

Uffizi Museum

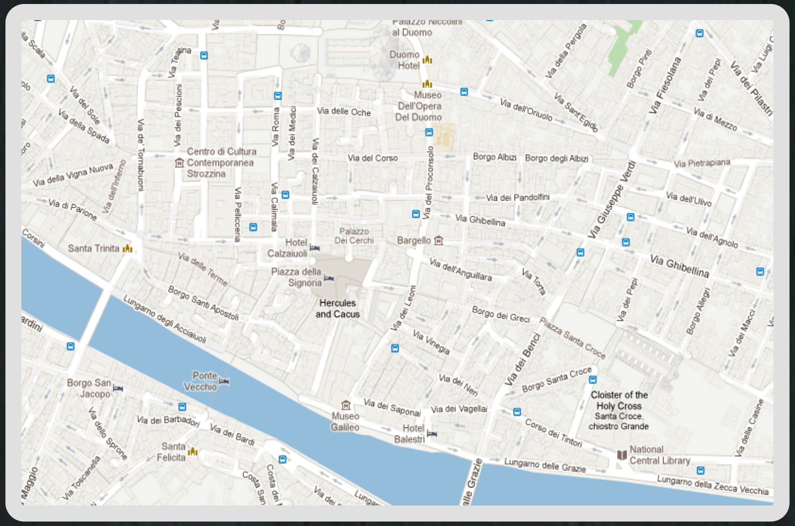
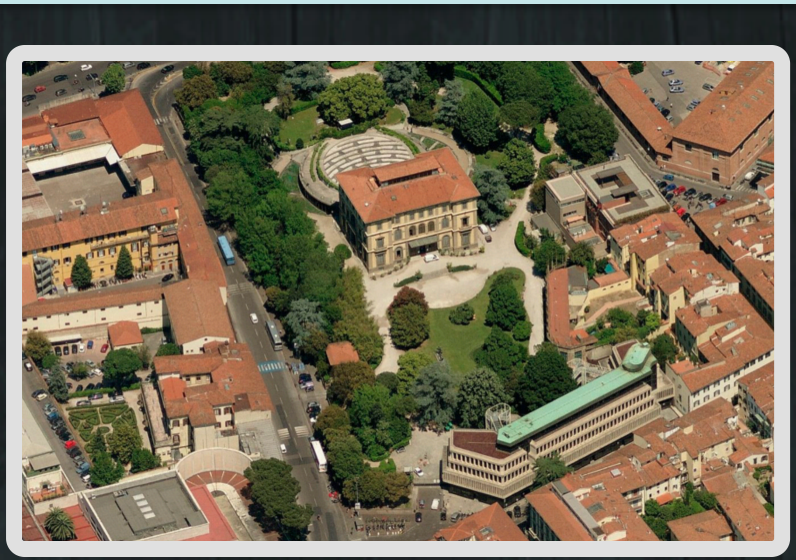
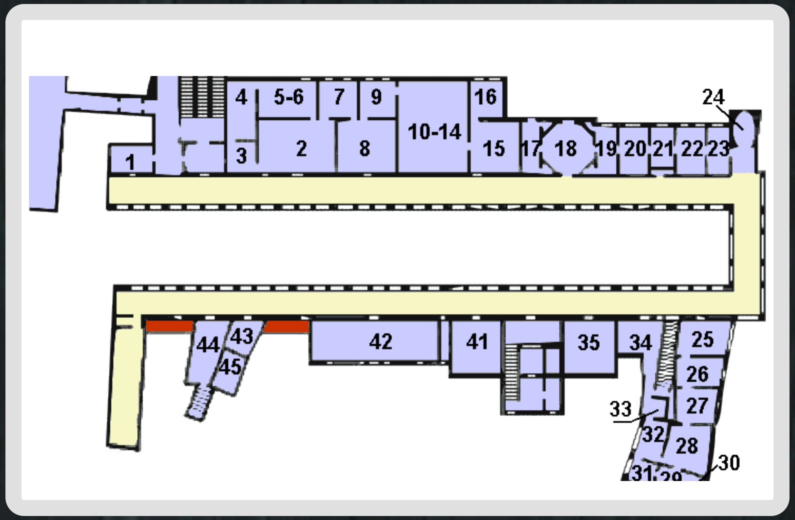


Uffizi Museum

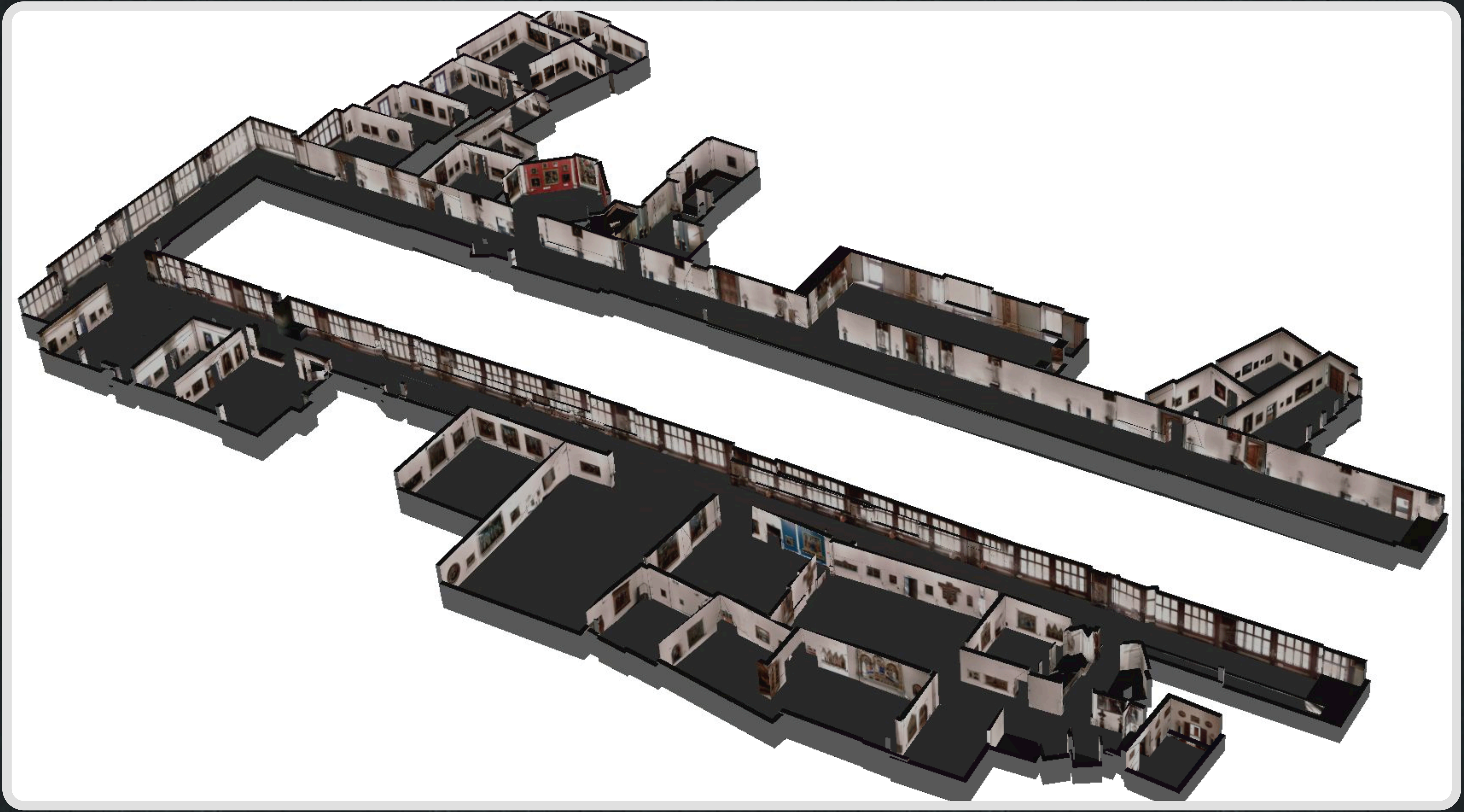


Where is "The Birth of Venus"?

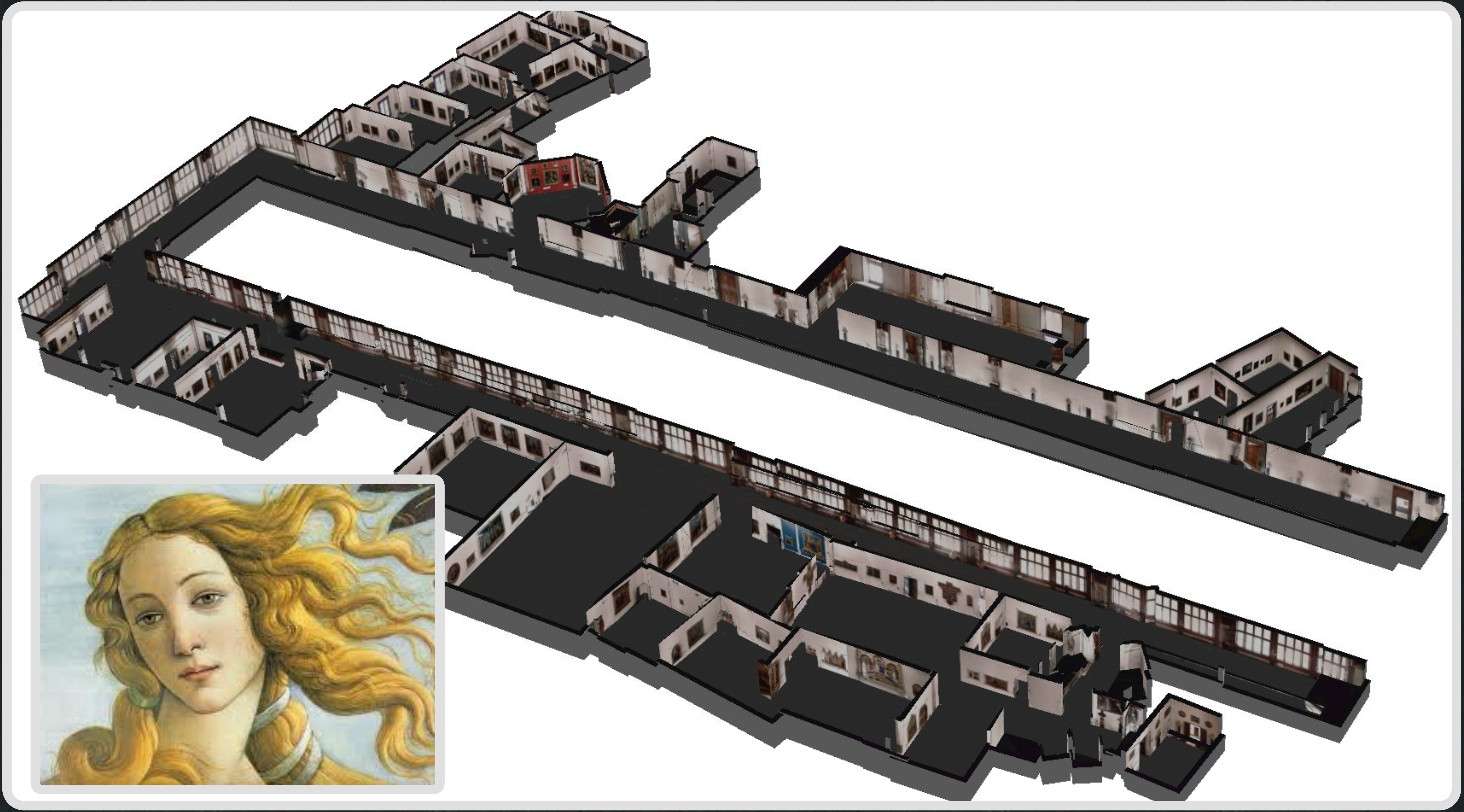
Cartography

	2D Line Drawing	3D Realistic Maps
Outdoor		
Indoor		?

Photorealistic Indoor Maps

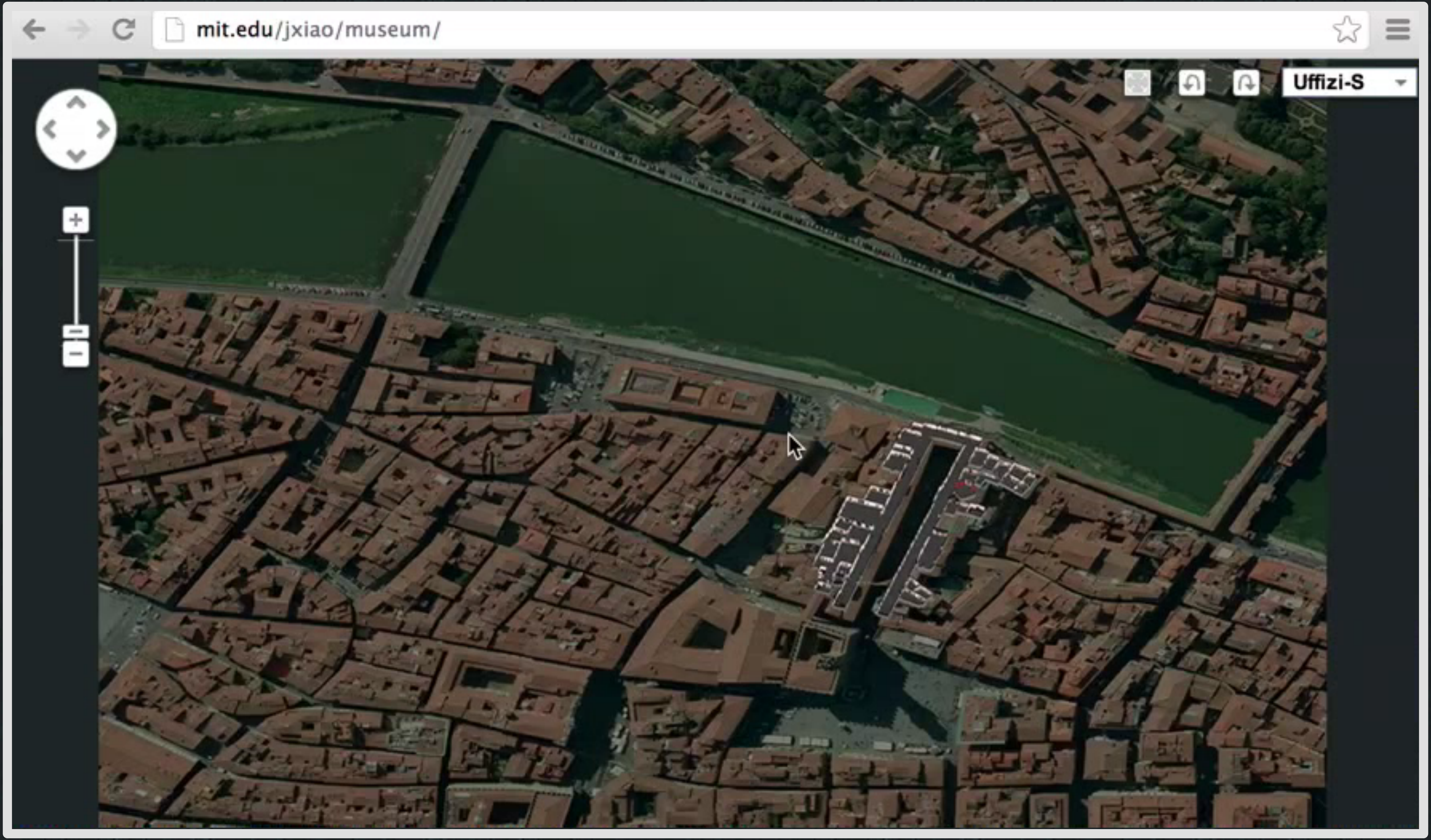


Photorealistic Indoor Maps



Where is "The Birth of Venus"?

Photorealistic Indoor Maps



Data-driven Brute-force Approach

Data-driven Brute-force Approach

SUN Database. Xiao et al. 2010.

Data-driven Brute-force Approach

2-step algorithm

Data-driven Brute-force Approach

1. Remove ceiling.
2. Take pictures from aerial viewpoints.

What is wrong?

What is wrong?

I don’t own the museums.

Old Approach

1. Remove ceiling.
2. Take pictures from aerial viewpoints.

New Approach

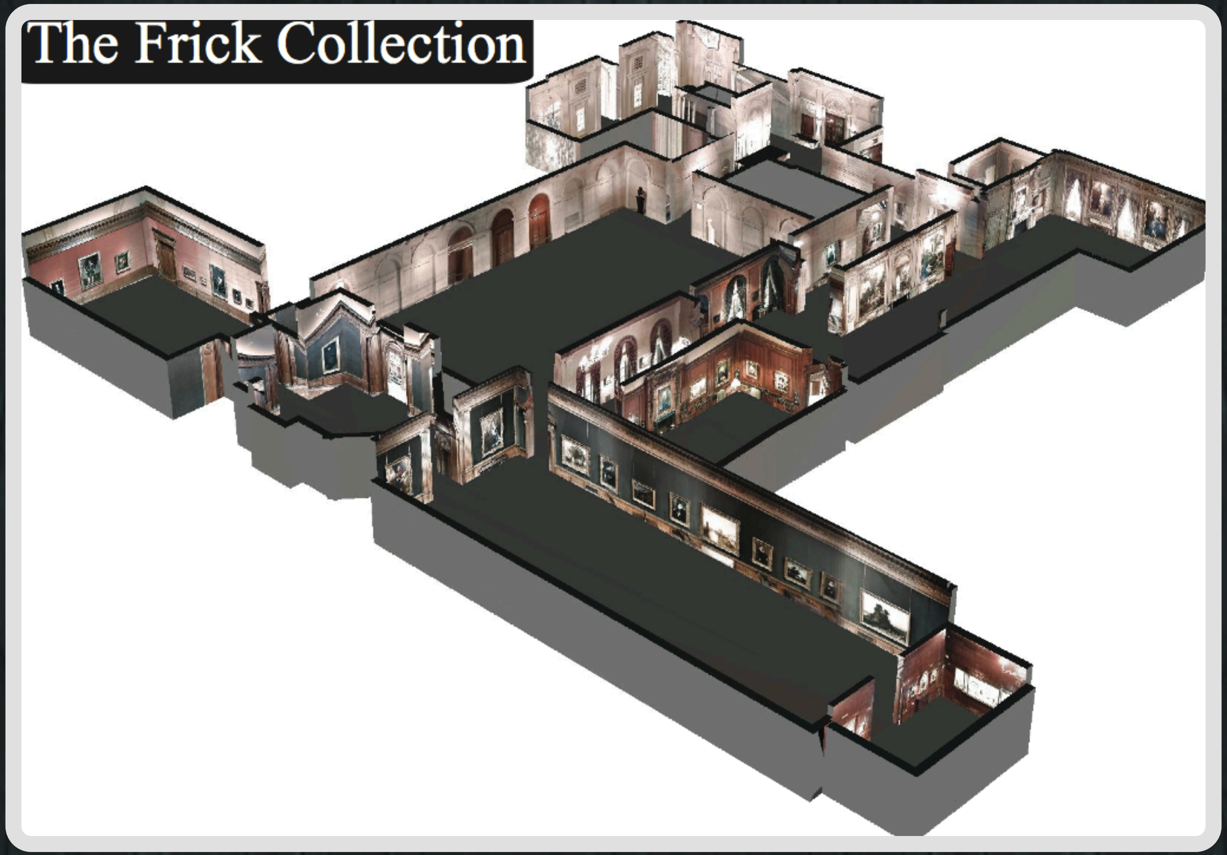
0. Own the museums.
1. Remove ceiling.
2. Take pictures from aerial viewpoints.

New Approach

0. Reconstruct the museums.
1. Remove ceiling.
2. Take pictures from aerial viewpoints.

Our Goal

- Global texture-mapped 3D model
- Optimize for aerial viewing
- Enable effective indoor navigation



Size matters!

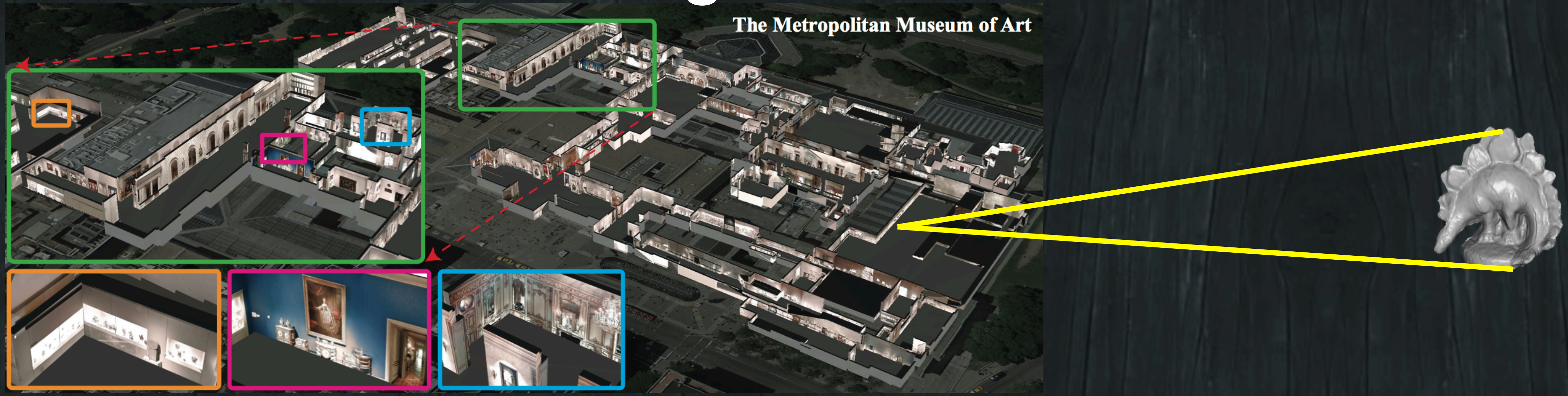
Existing methods

[Xue 11], [Hongxing 10], [Kolev 10],
[Li 10], [Song 10], [Xi 10], [Xiao 09], [Furukawa 09],
[Jancosek 09], [Vu 09], [Xiao 08], [Bradley 08],
[Furukawa 08], [Zach 08], [Xiao 07] , [Gargallo 07],
[Goeselle 07], [Habbecke 07], [Merrell 07],
[Sinha 07], [Sormann 07], [Stark 07], [Vogiatzis 07],
[Zach 07], [Zaharescu 07], [Furukawa 06],
[Goeselle 06], [Hornung 06], [Tran 06],
[Strecha 06], [Pons 05], [Vogiatzis 05],
[Hernandez 04], [Kolmogorov 02], ...

- Require accurate calibration
- Produce a “dense” model
  - Susceptible to errors
  - Not scalable
- Mostly small scale

22,610 triangles

> 1,000,000 triangles

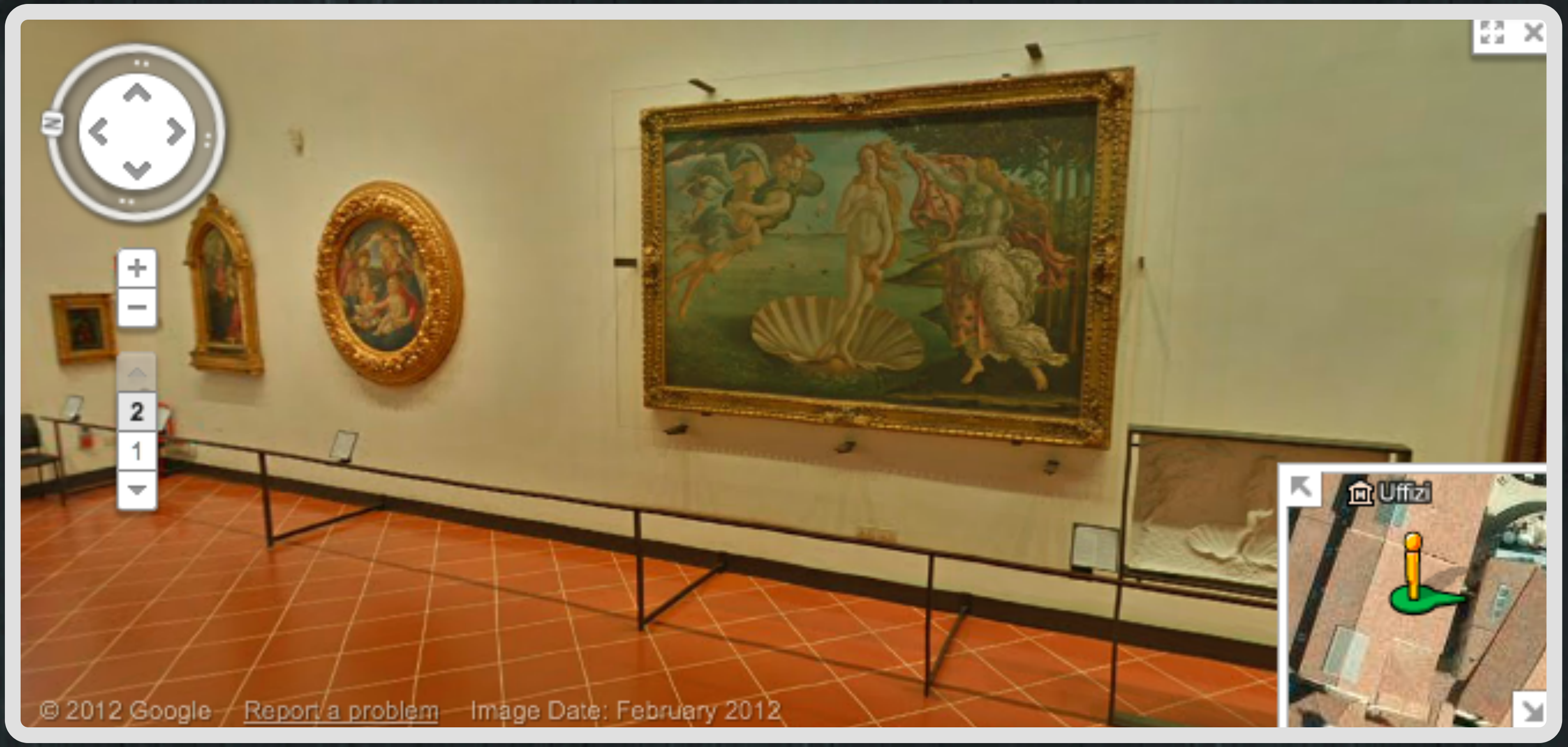


System Pipeline

1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints

System Pipeline

1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints

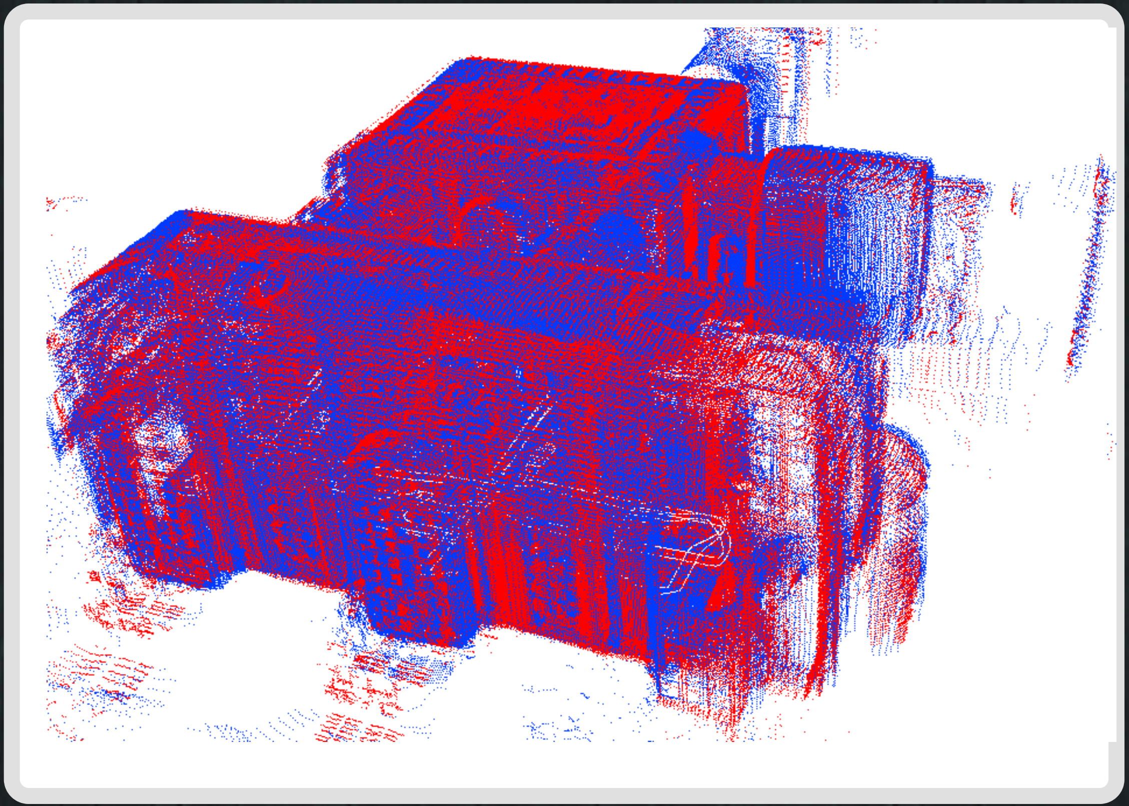


Art Project

www.GoogleArtProject.com

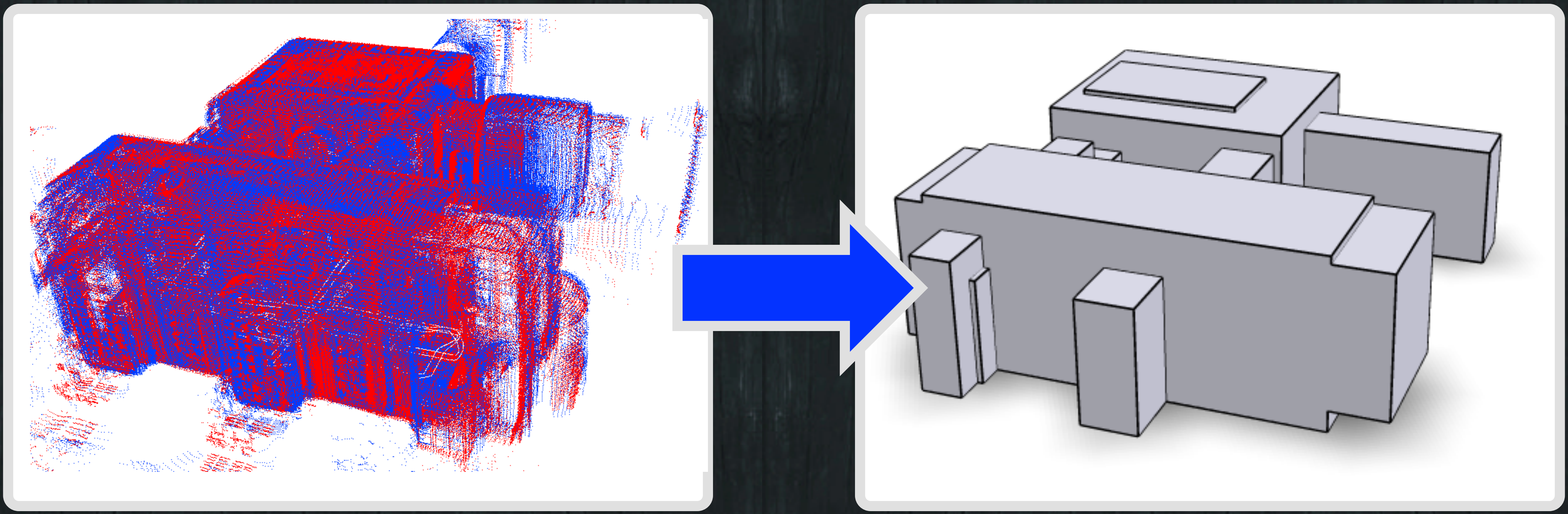
System Pipeline

1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints



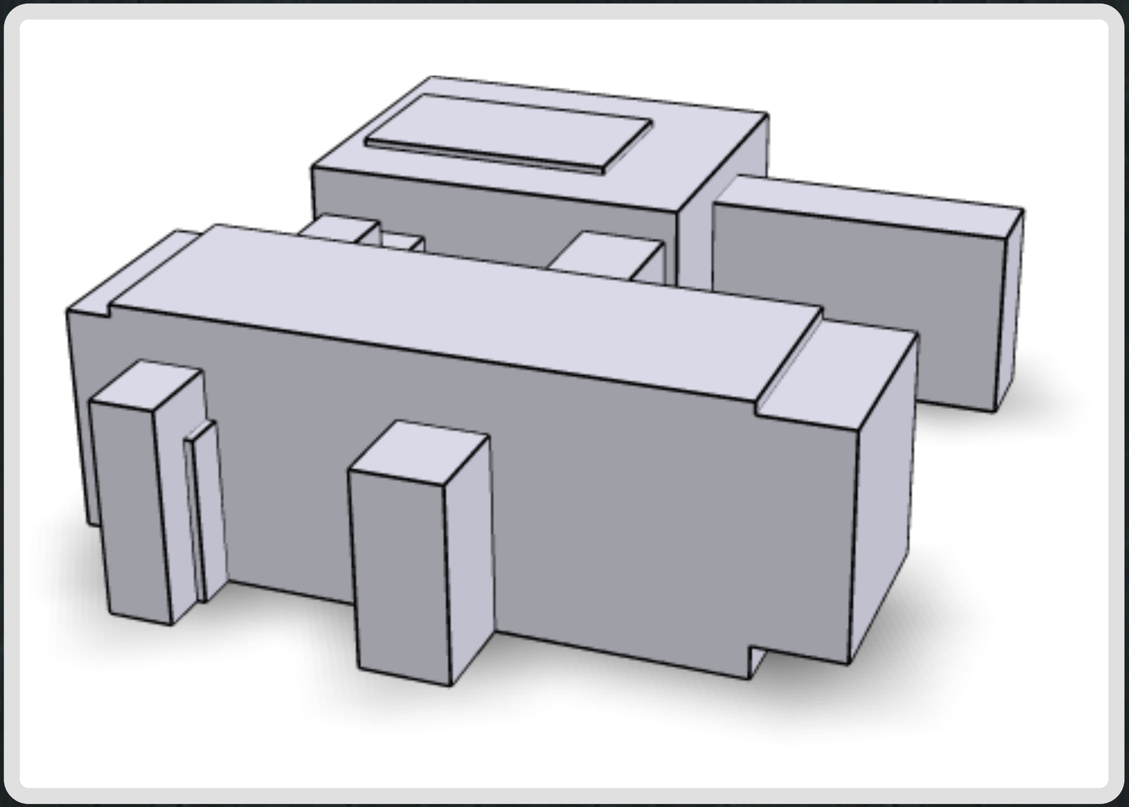
System Pipeline

1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints



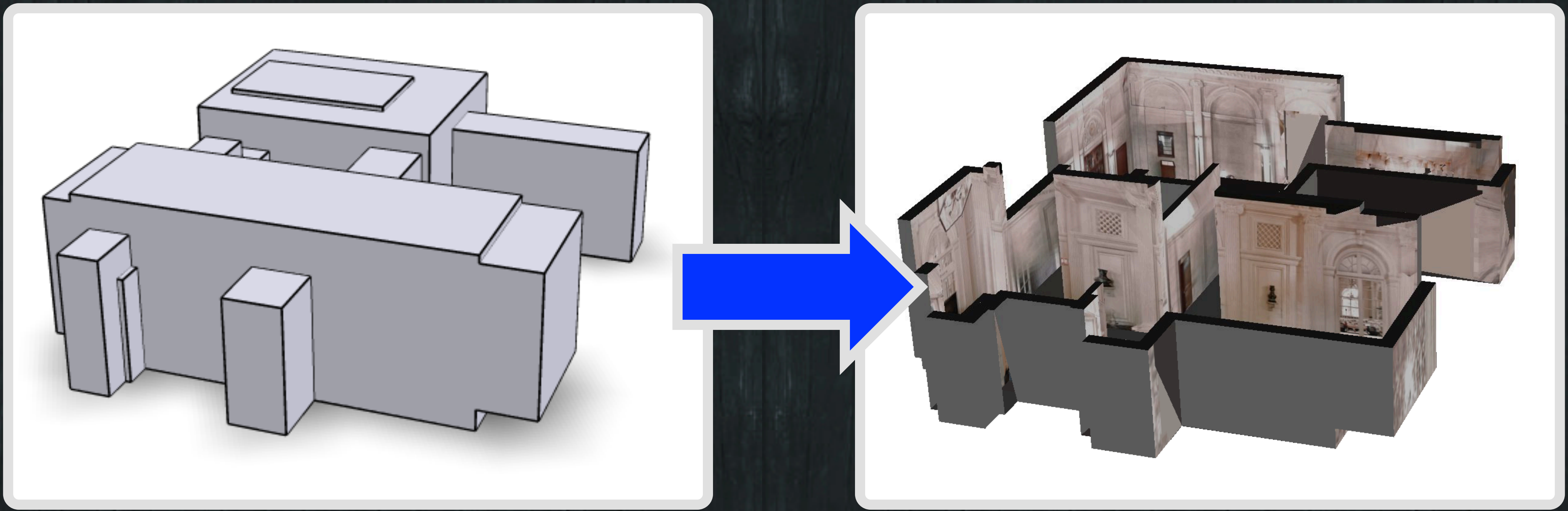
System Pipeline

1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints

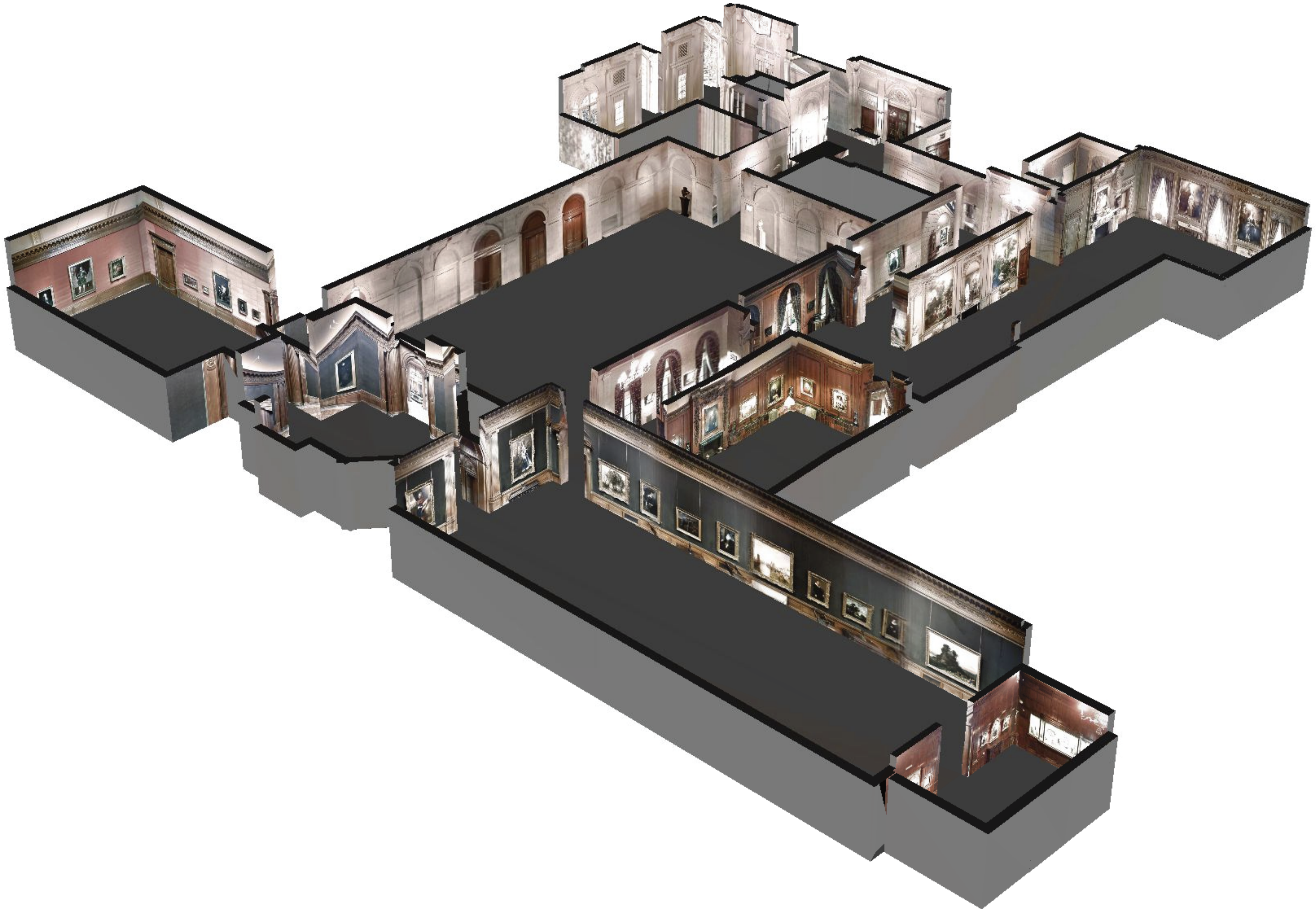


System Pipeline

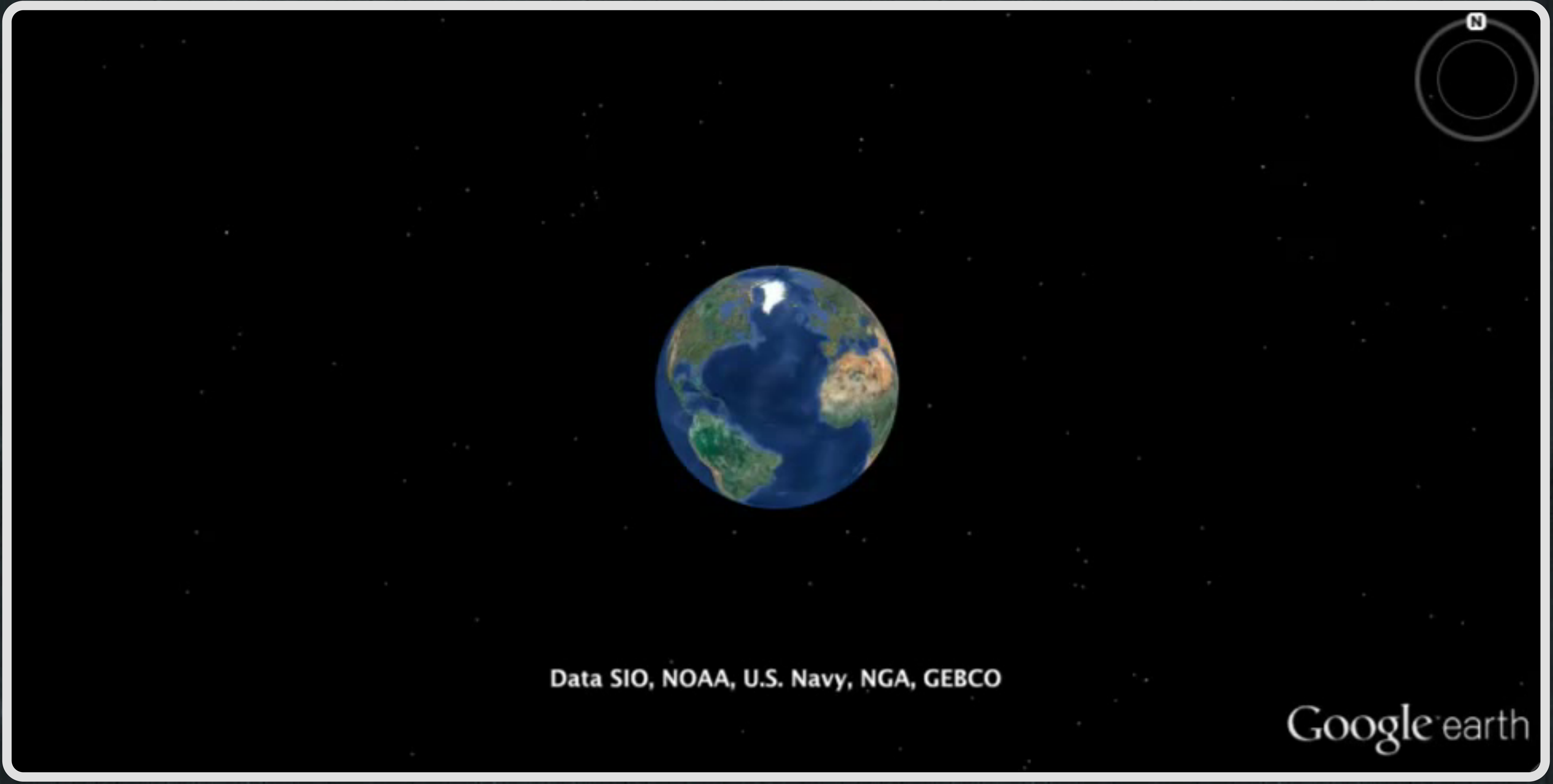
1. Take pictures inside the rooms
2. Reconstruct the 3D shape
3. Render from aerial viewpoints



Results



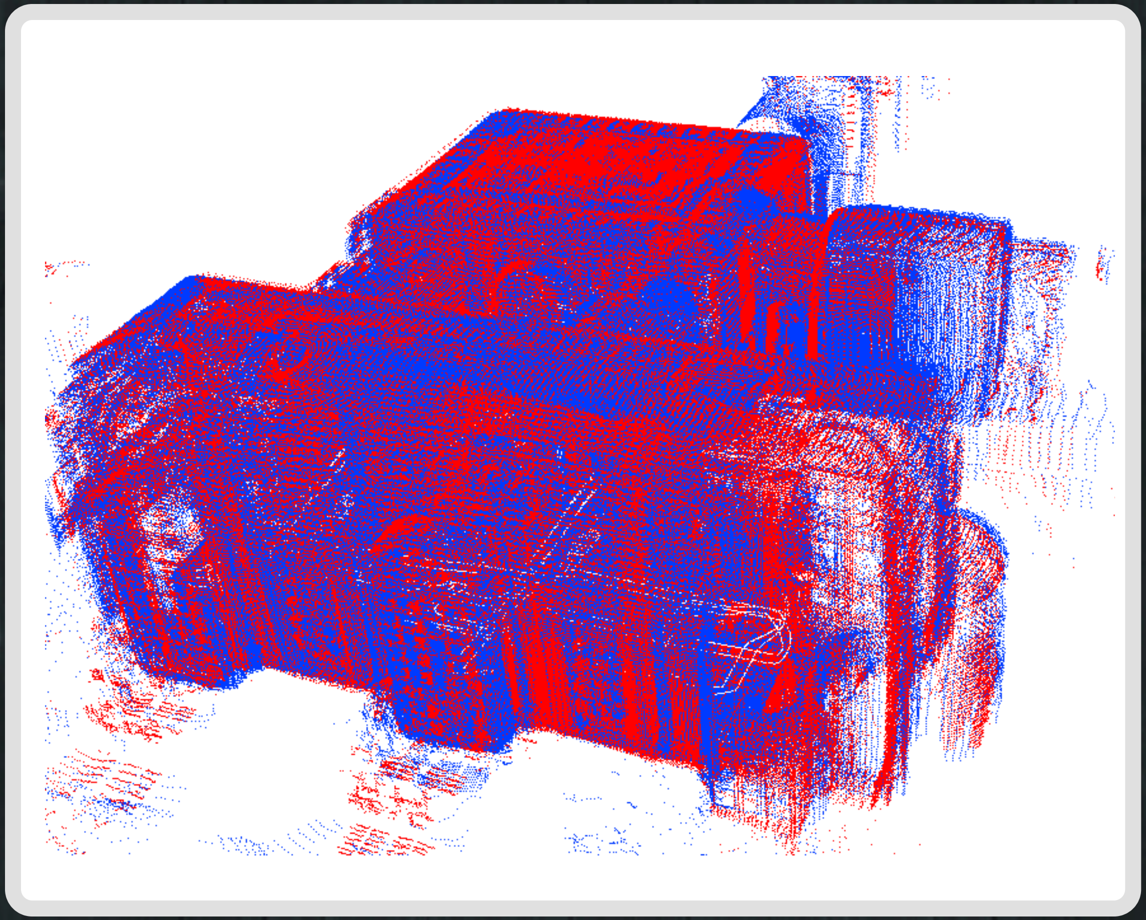
Effective Navigation



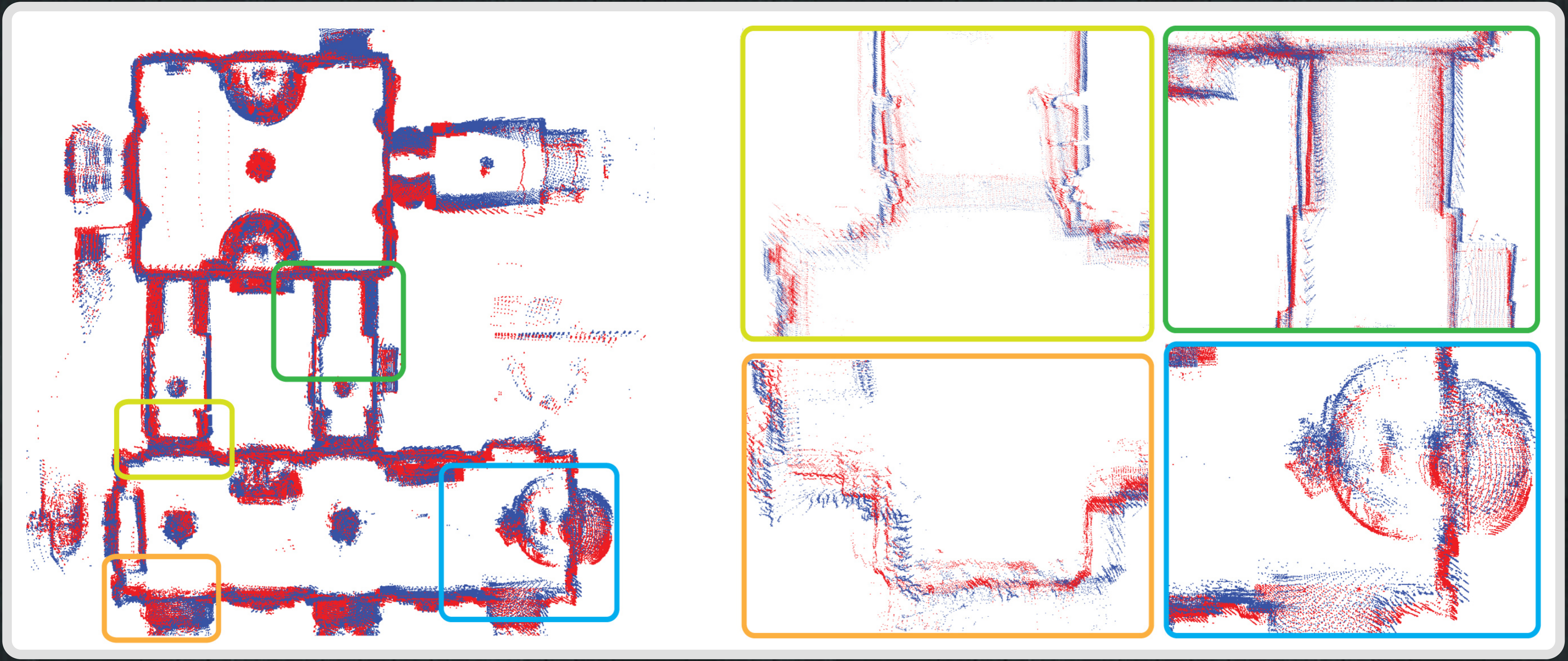
Reconstruction

Input

- Images (> 40,000)
- Laser points (> 200,000,000)
- Input collected over multiple sessions

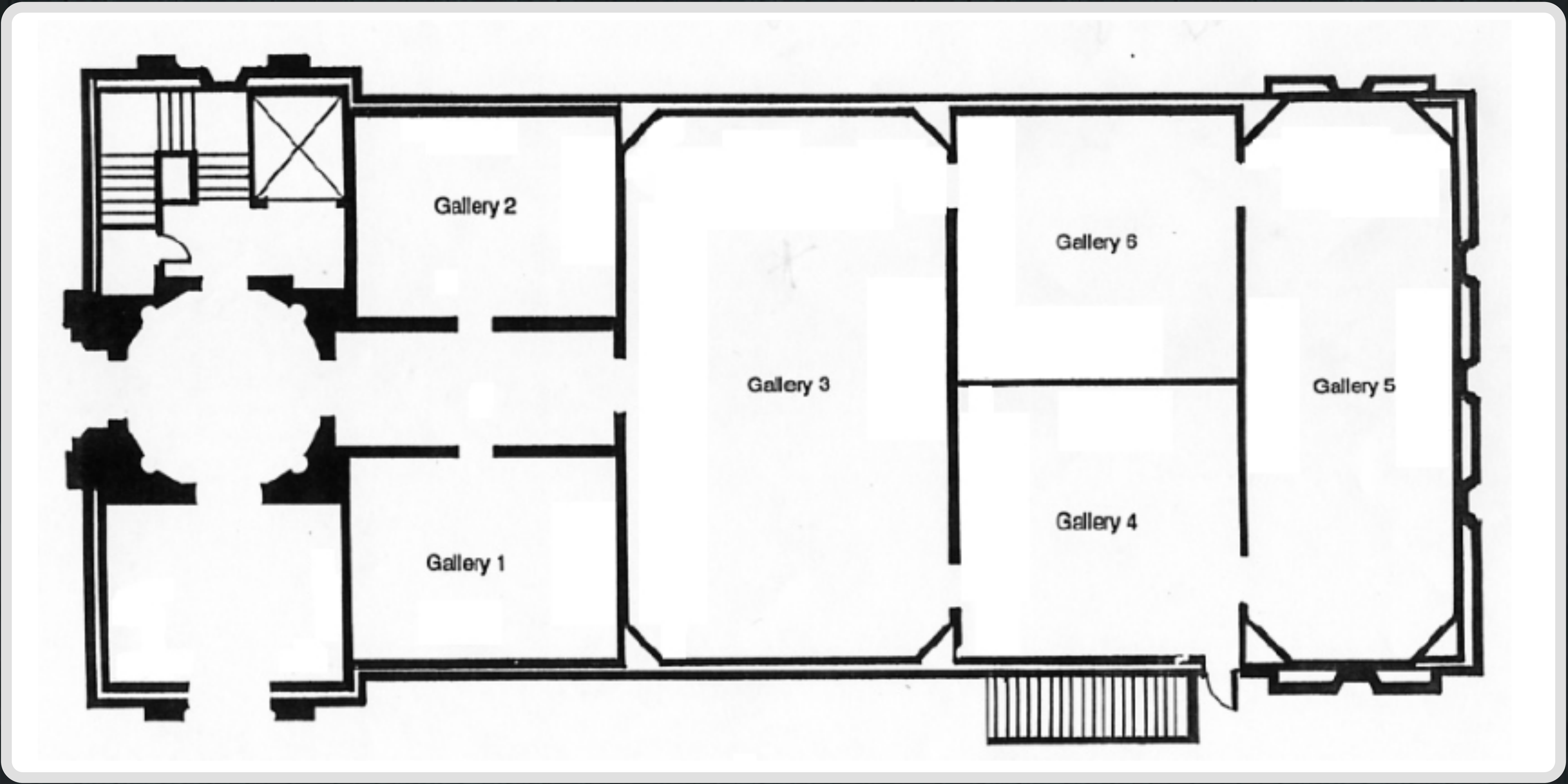


Noisy Laser Points



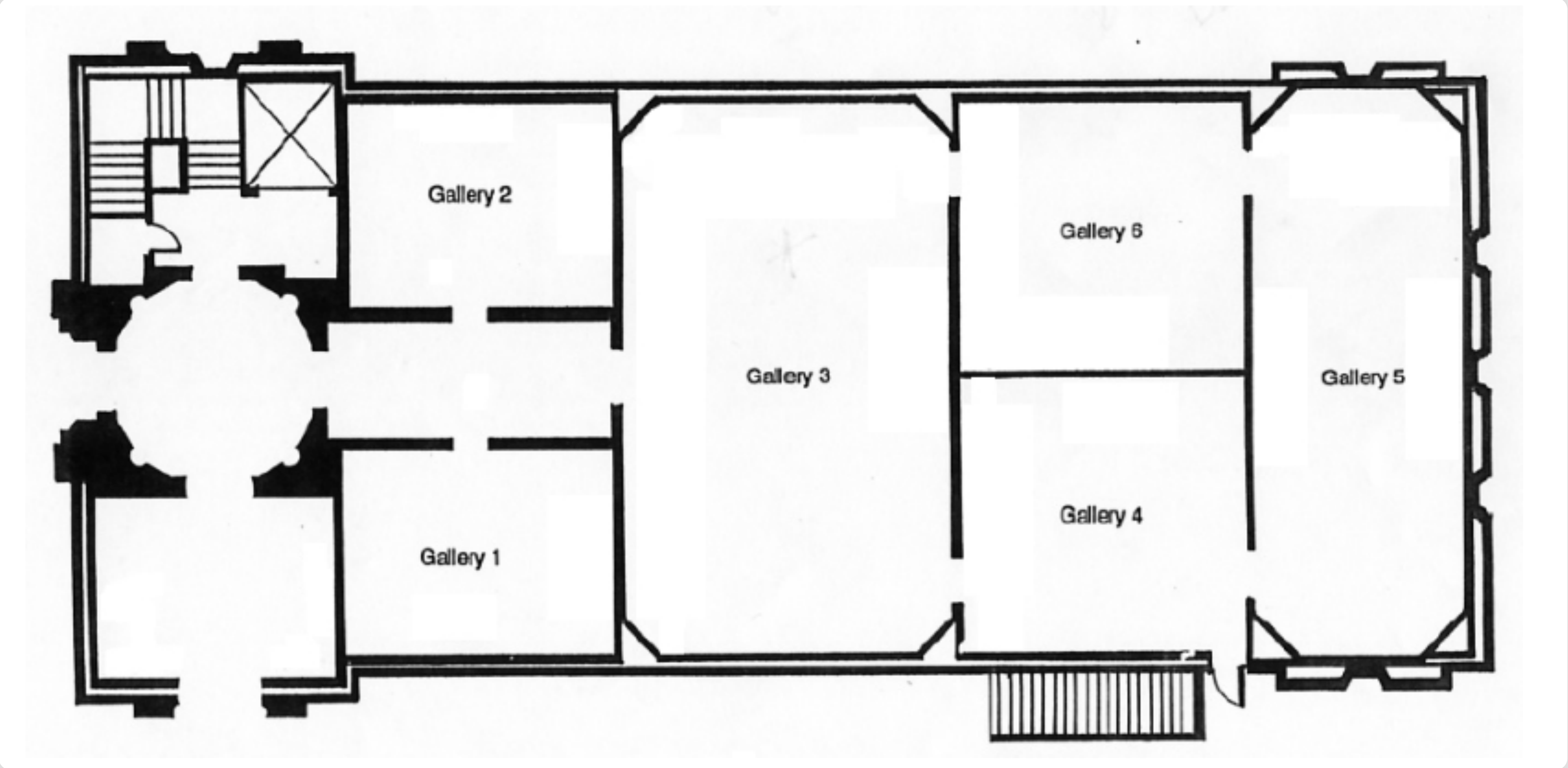
top-down view of input laser points
(two different colors represent two vertical laser range sensors)

Challenges for Indoor Scenes



Prevalence of thin structures

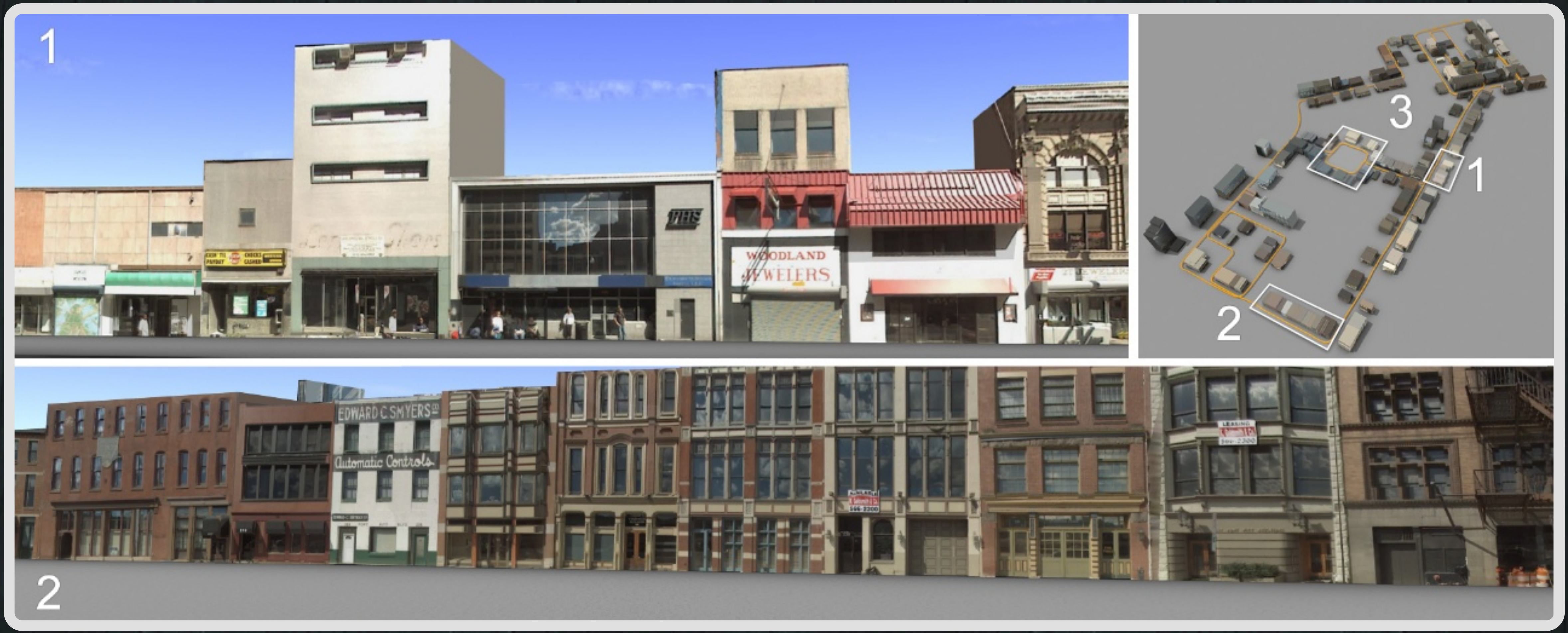
Opportunities for Indoor Scenes



Structural regularities
(planarity/orthogonality)

Exploit Structural Regularity

Image-based Street-side City Modeling, Xiao et al, 2009.



Xiao, Fang, Tan, Zhao, Ofek and Quan, 2008

Furukawa, Curless, Seitz and Szeliski 2009

Liu, Carlberg, Chen, Chen Kua and Zakhor, 2010

Adan and Huber, 2011

Gallup, Frahm and Pollefeys, 2010

Sinha, Steedly and Szeliski, 2009

Birchfield and Tomasi, 1999

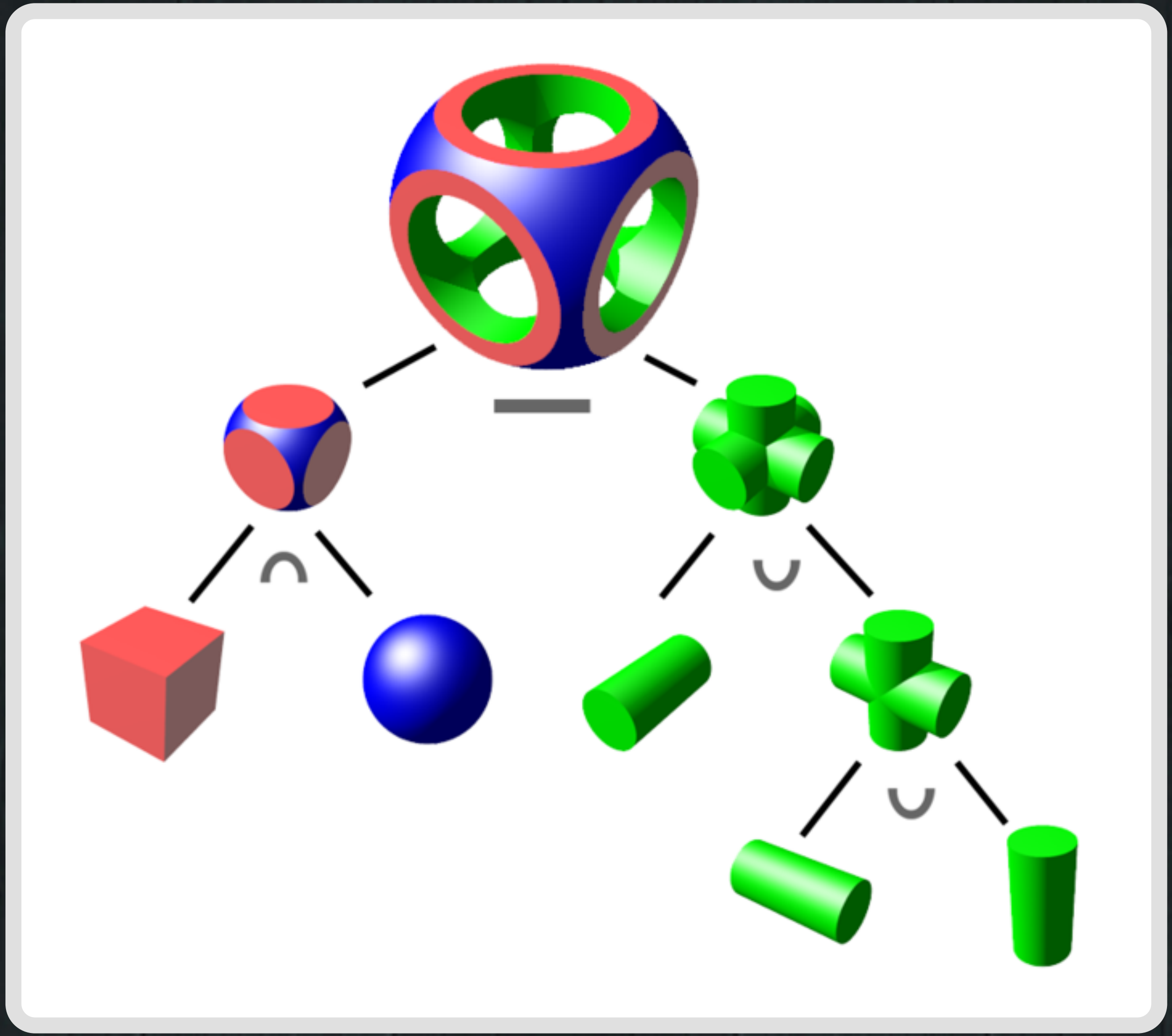
Wang and Adelson, 1994 …

Summary on Previous Work

- Structure prior enforced in 2D (depth or façade)
- Visualization limited to ground level
- Assume near-perfect calibration

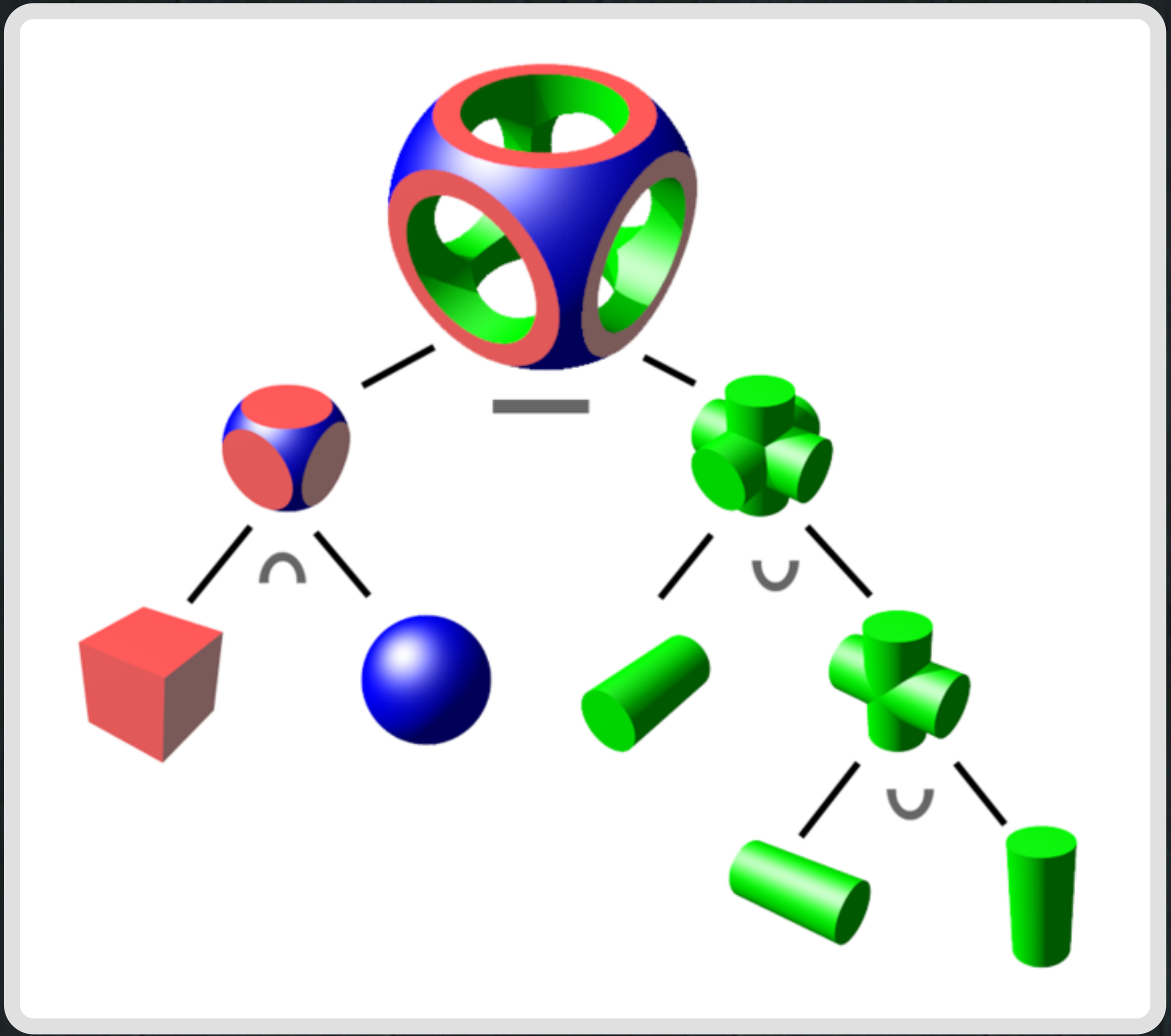
InverseCSG Algorithm

Constructive Solid Geometry (CSG)

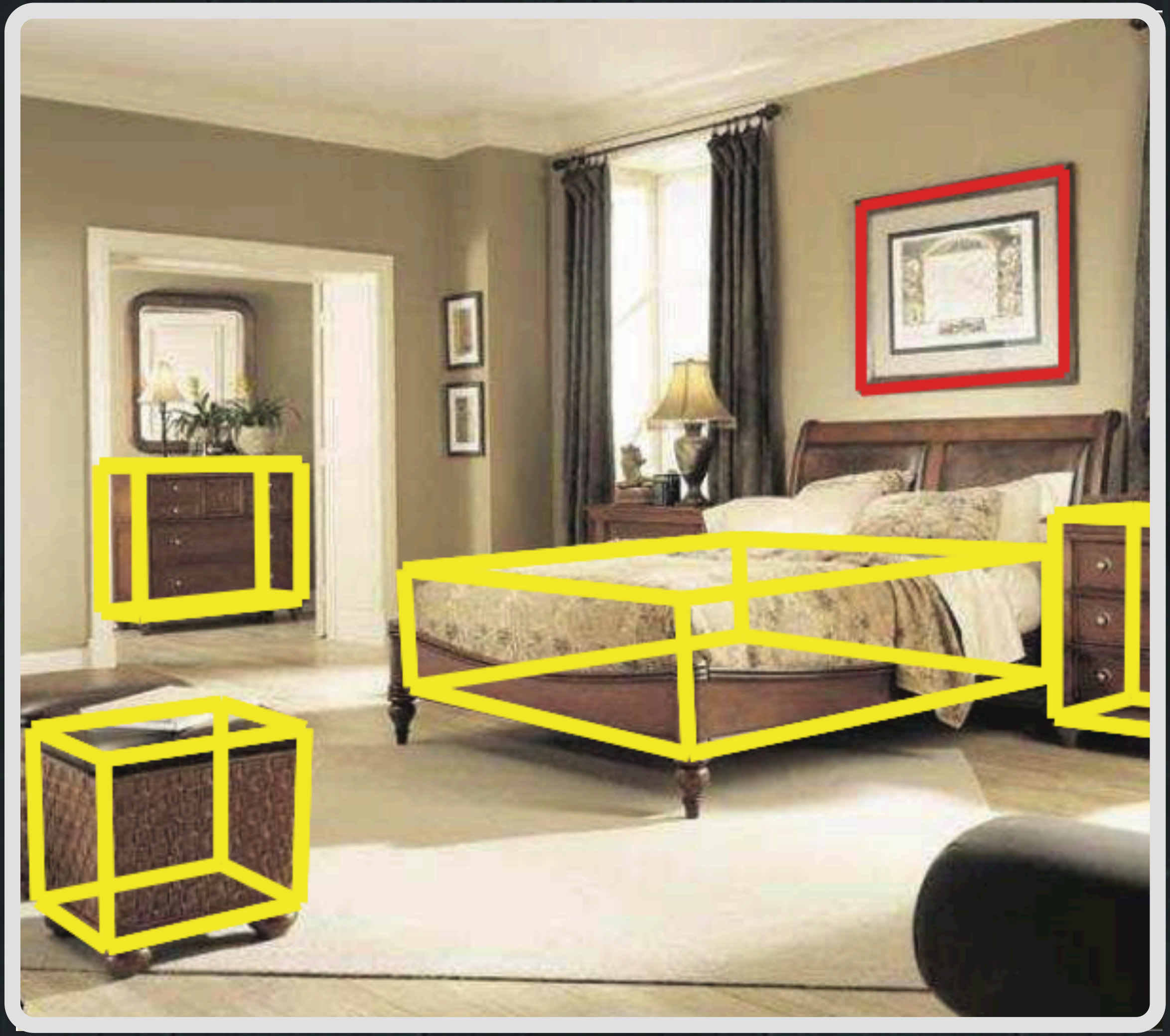


InverseCSG Algorithm

Constructive Solid Geometry (CSG)



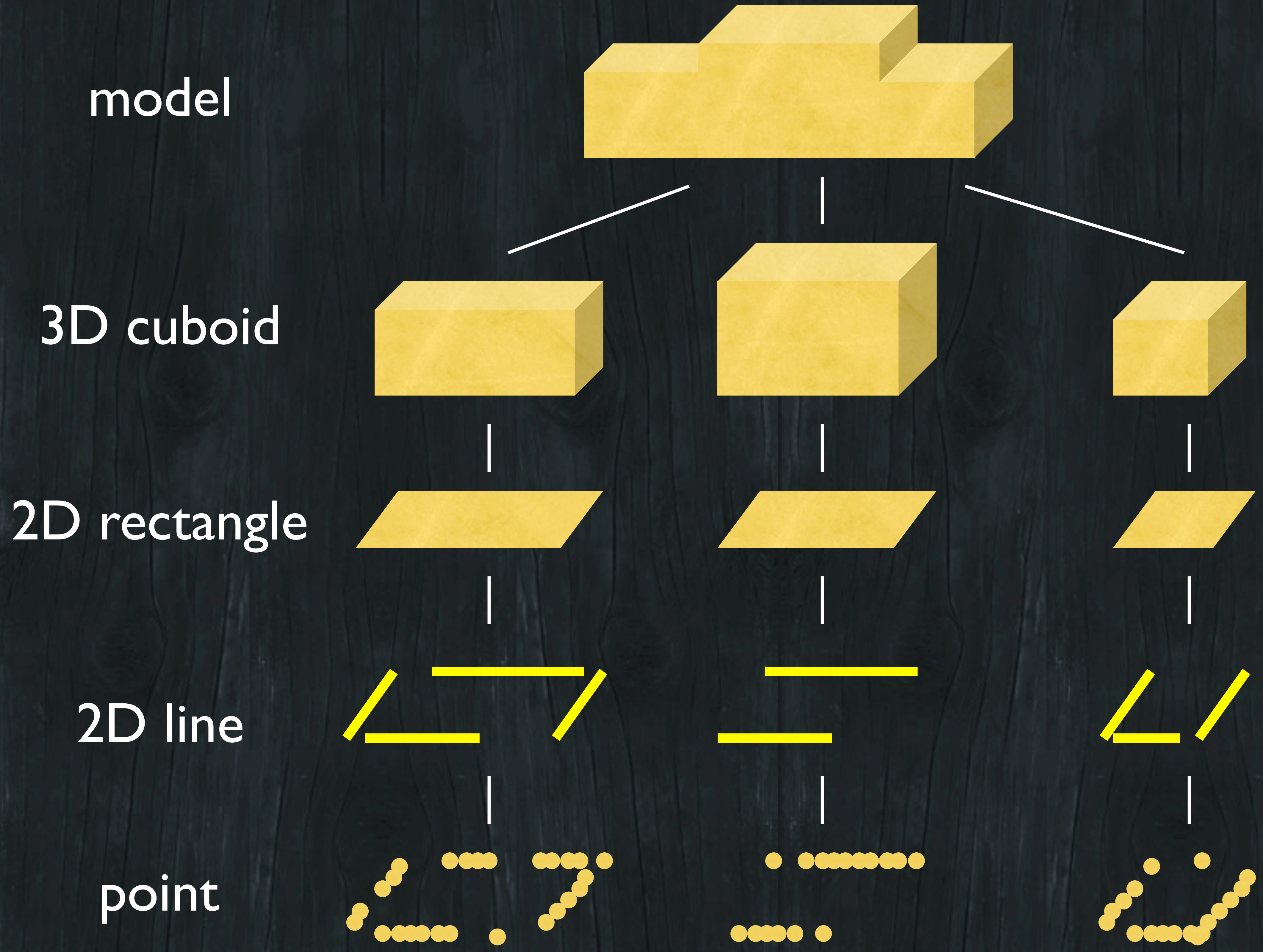
Cuboids as primitives



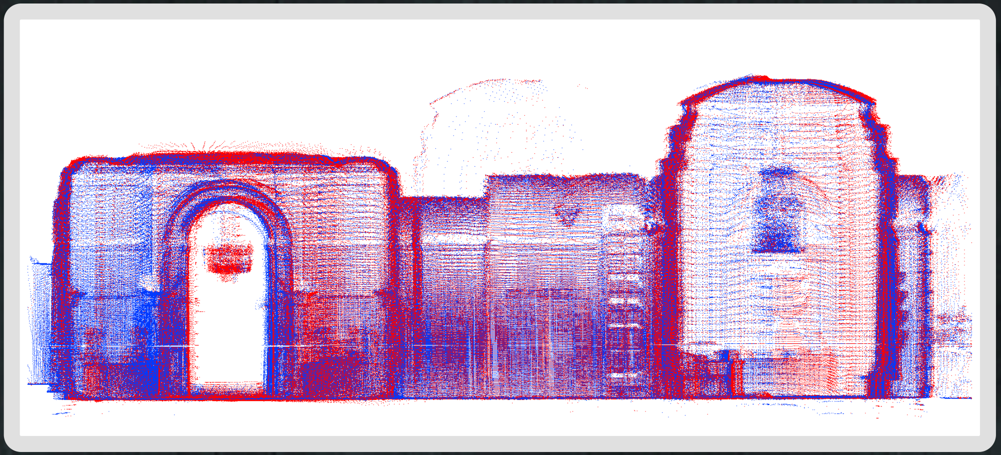
Xiao et al. NIPS 2012

Xiao et al. Siggraph Asia 2012a

Bottom-up & Top-down

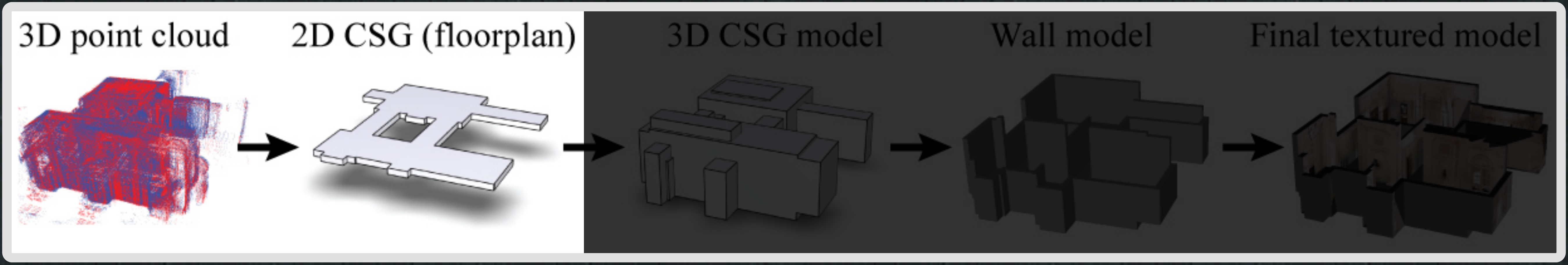


Cut into Slices

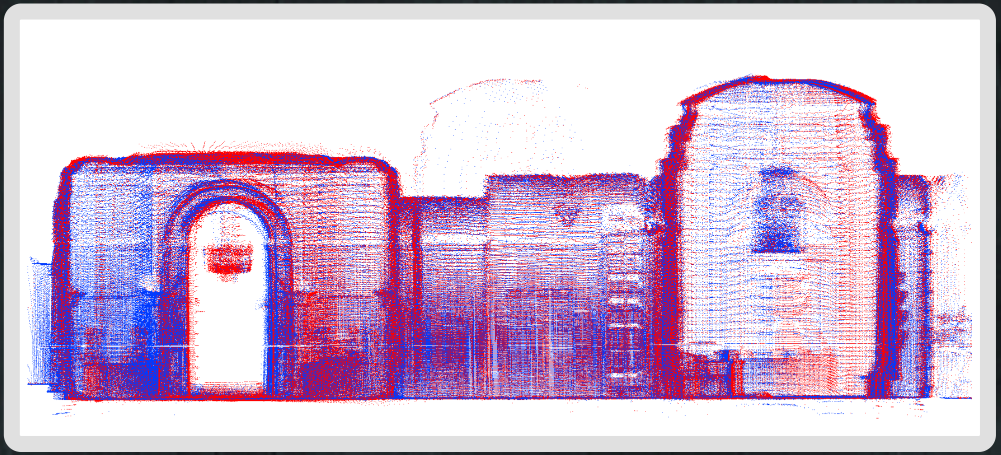


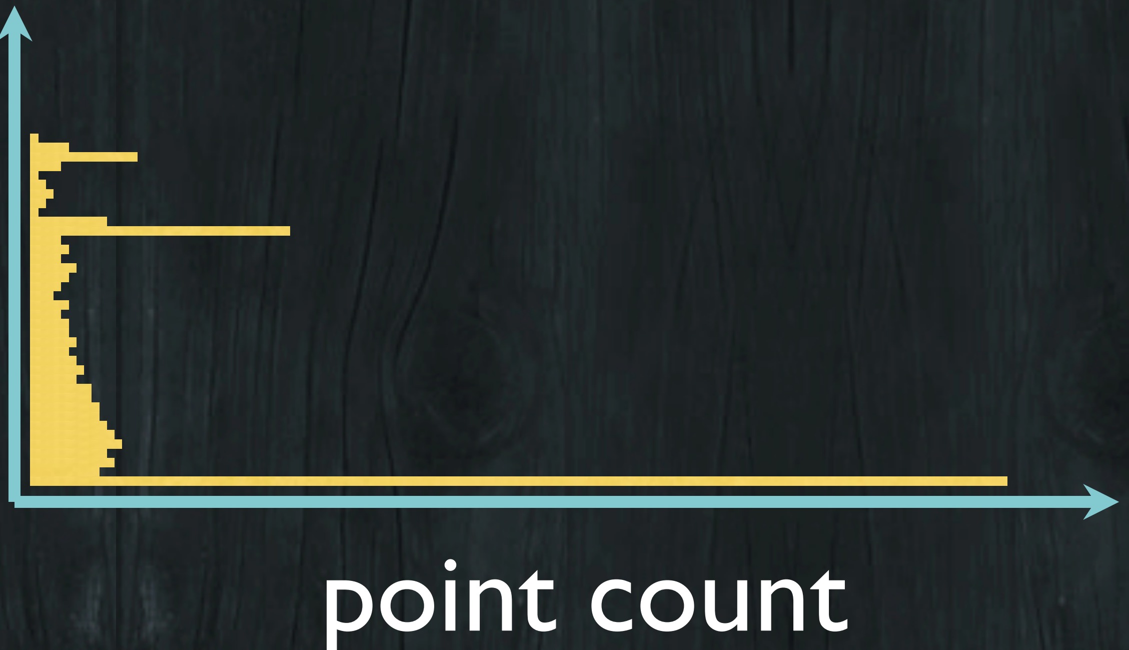
side view

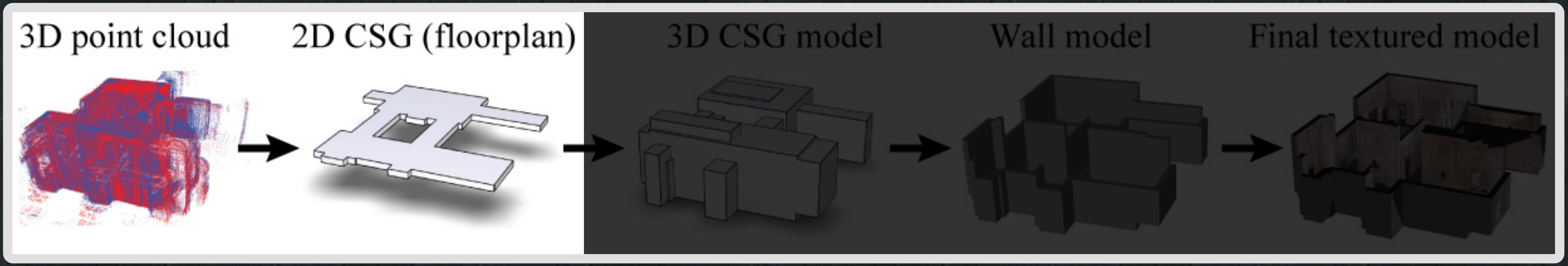
gravity



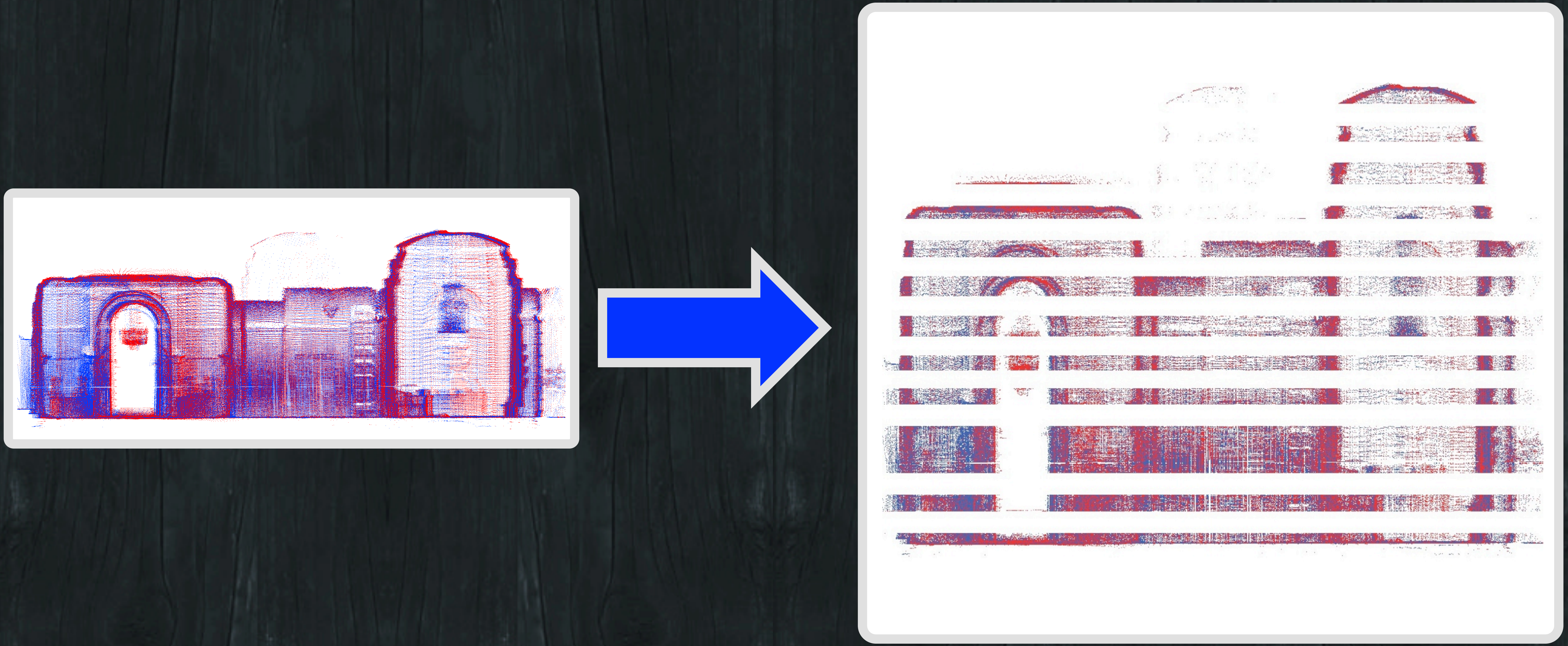
Cut into Slices

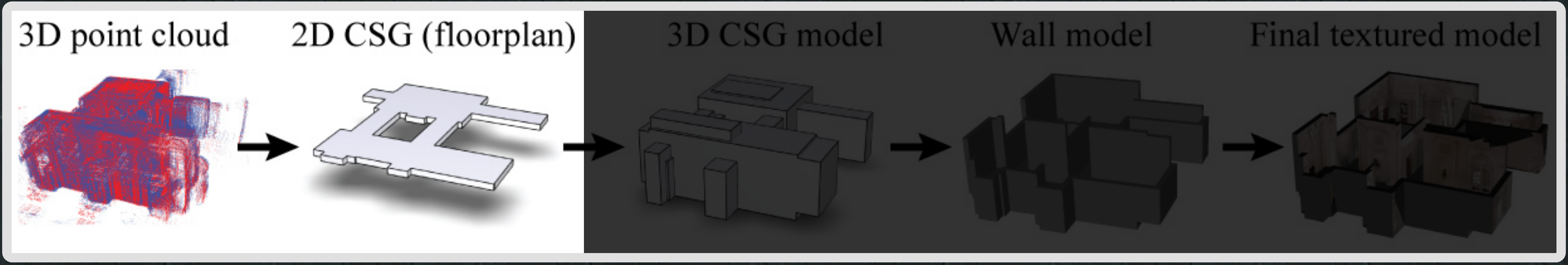






Cut into Slices

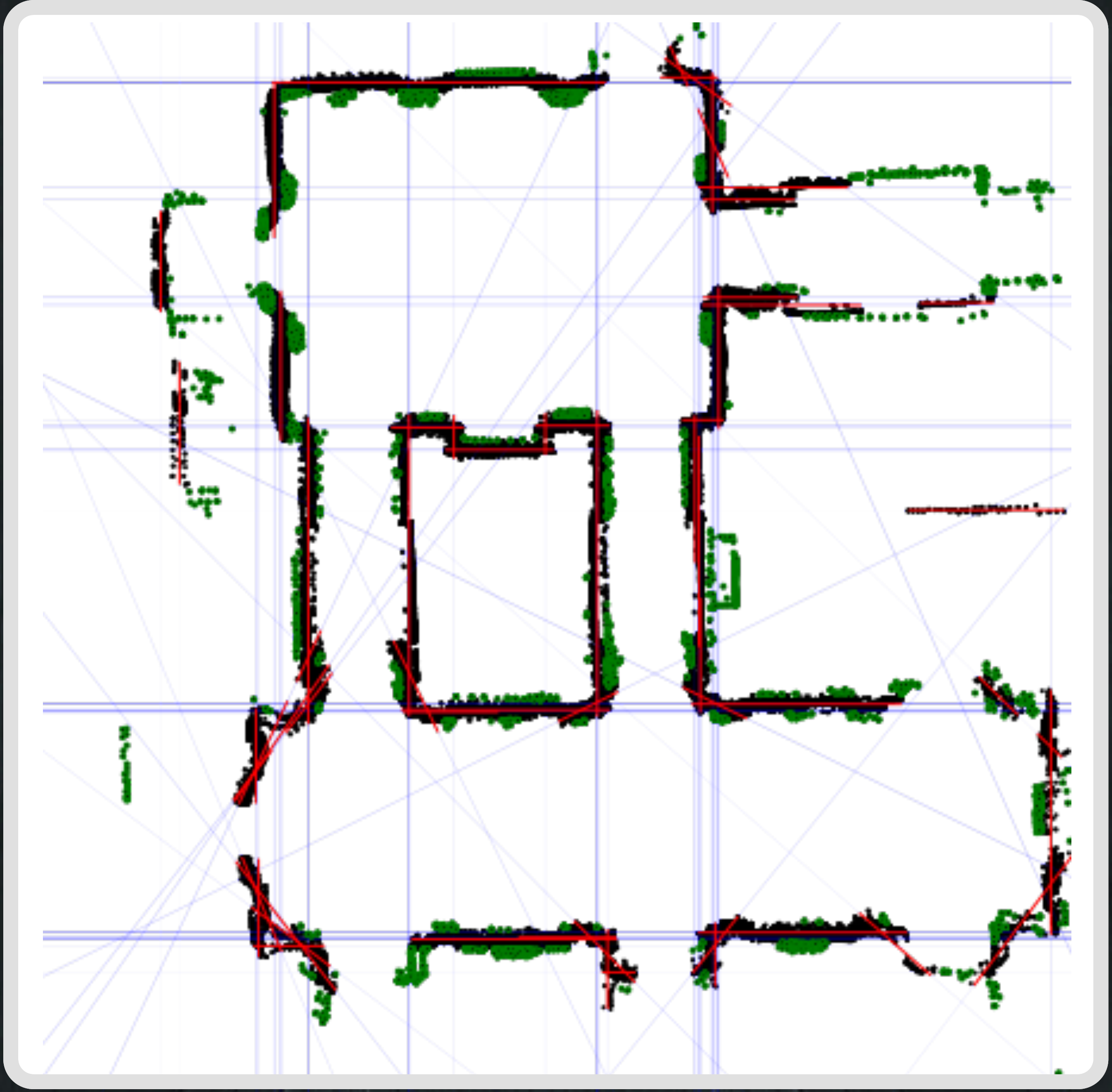


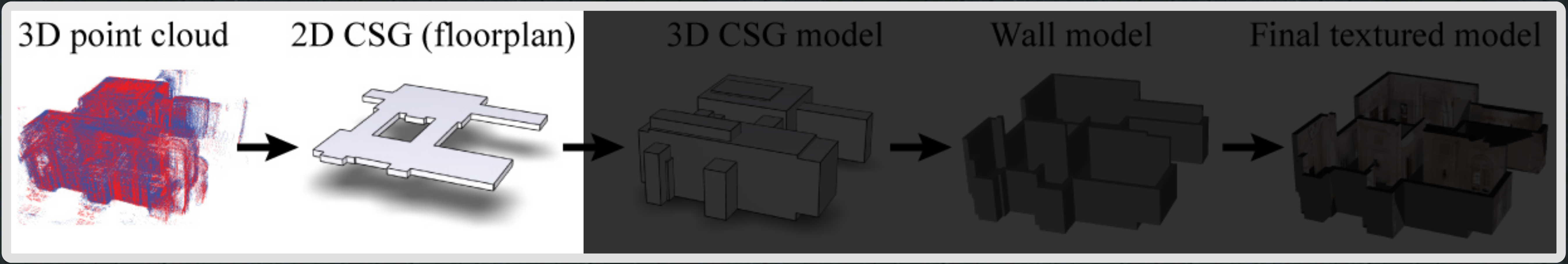


2D CSG Reconstruction

1. Generate primitives

2. Choose a subset

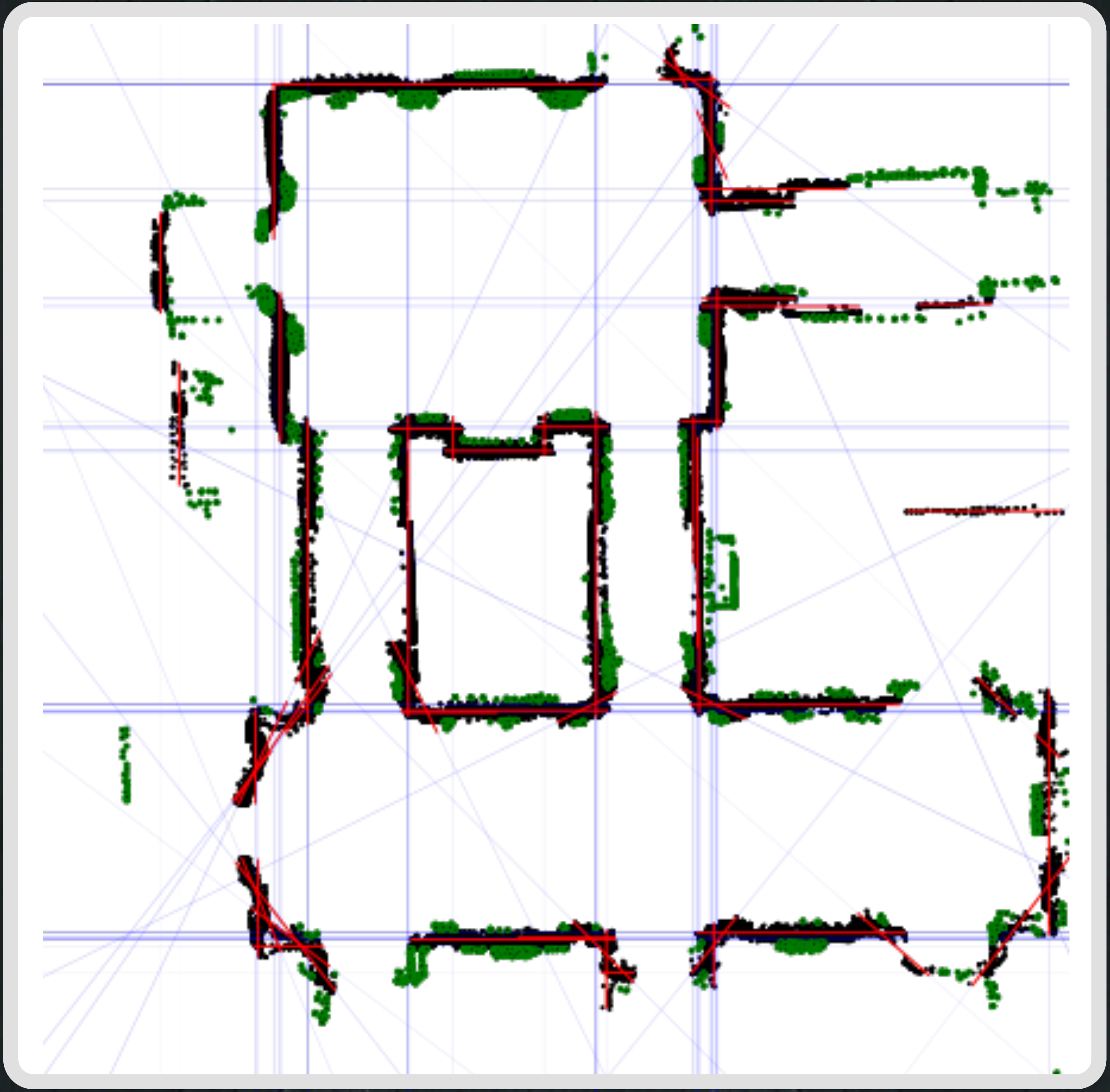




2D CSG Reconstruction

1. Generate primitives

point → line



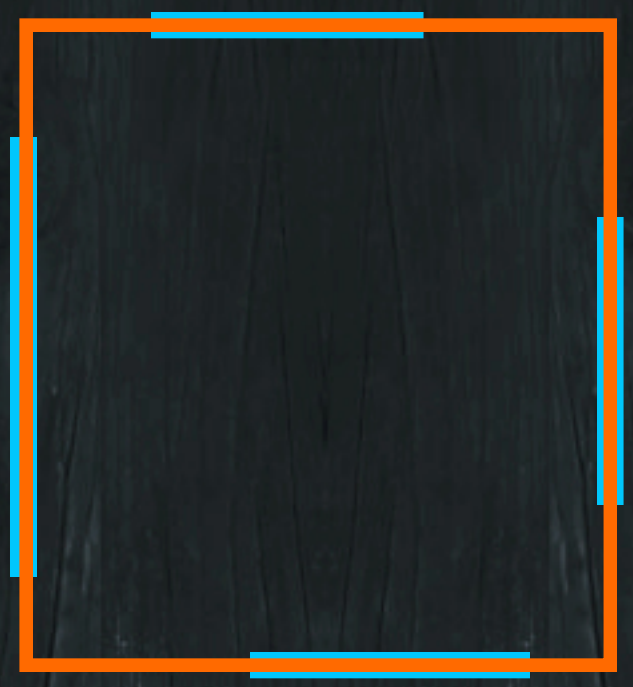
2D CSG Reconstruction

1. Generate primitives

point → line

line → rectangle

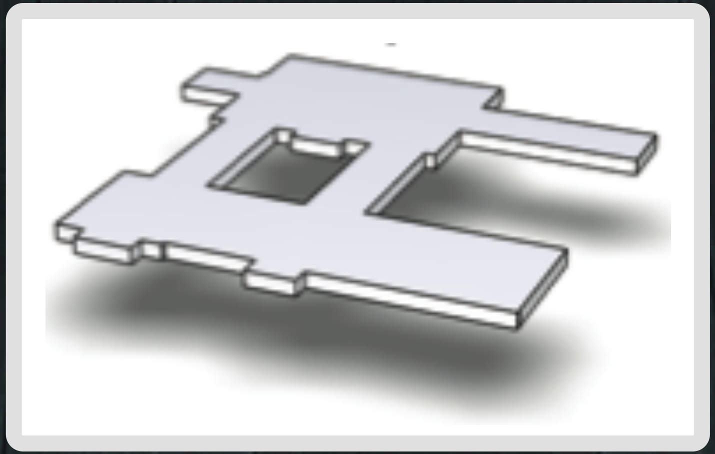
From 4 line segments

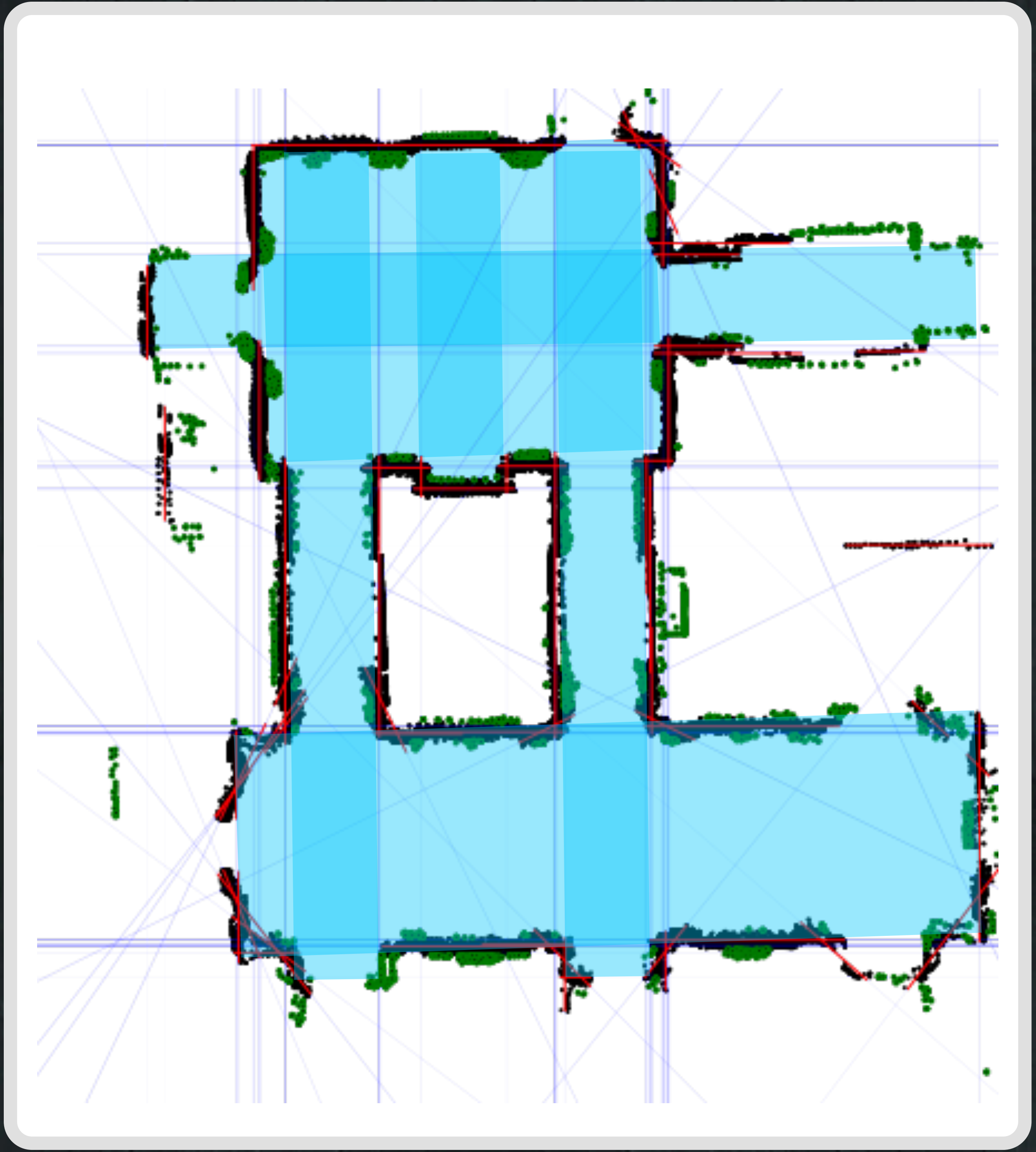


2D CSG Reconstruction

1. Generate primitives

2. Choose a subset



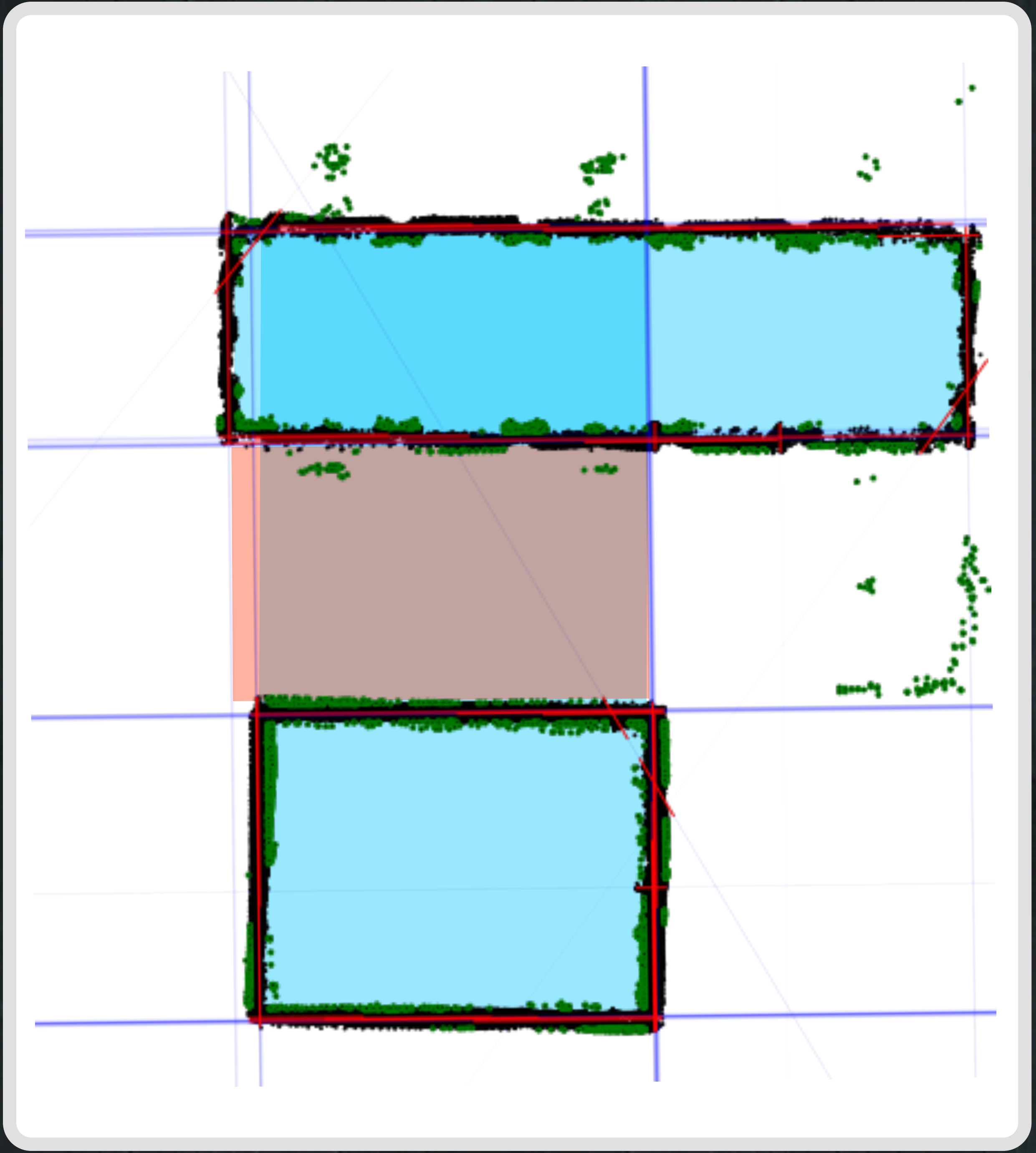


2D CSG Reconstruction

1. Generate primitives

2. Choose a subset

Repeat in each slice



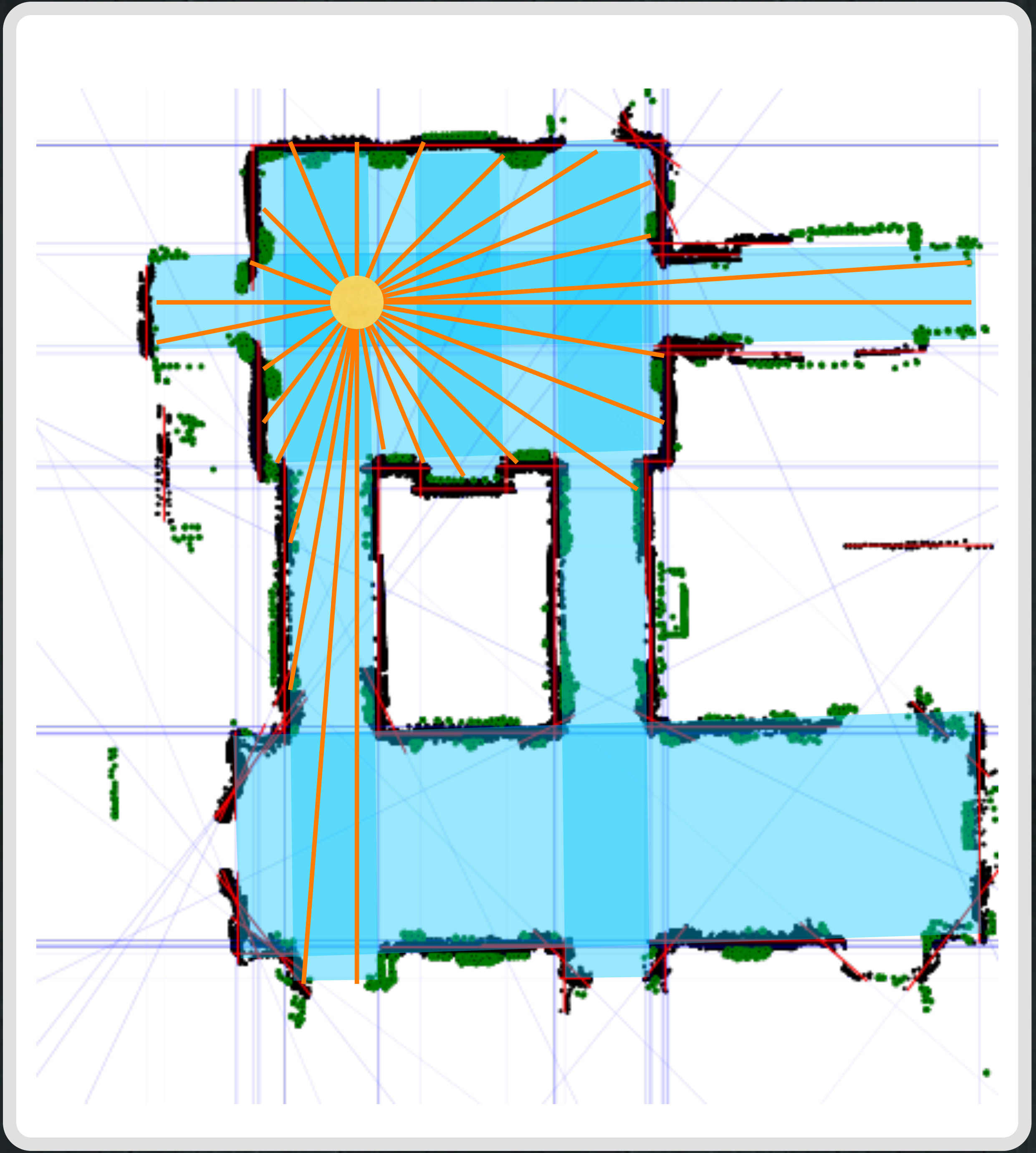
2D CSG Reconstruction

Explain the data

- Free space
- Laser points

Simple

- Regularization



Objective Function

$$\underline{E_1(T)} = \frac{\{\text{Sum of } \boxed{\text{free-space}} \text{ scores inside } T\}}{\{\text{Total sum in the domain without negative scores}\}}$$

$$\underline{E_2(T)} = \frac{\{\text{\# of } \boxed{\text{points on the surface}} \text{ of } T\}}{\{\text{total \# of points}\}}$$

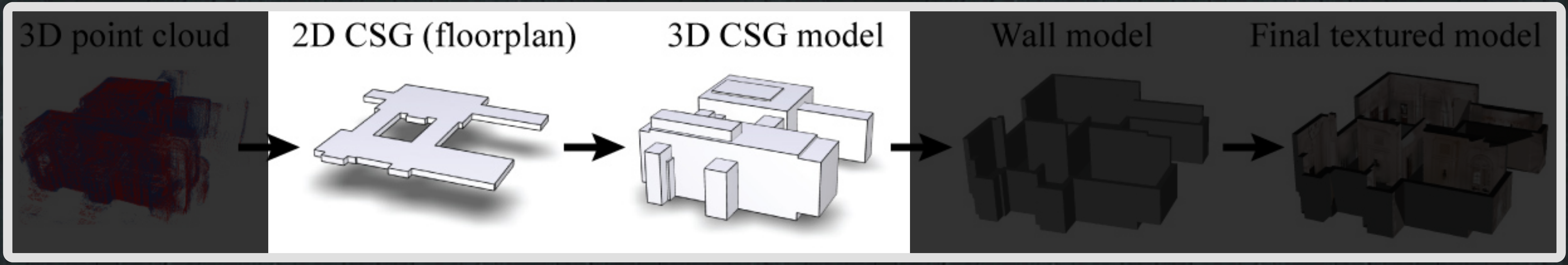
$$\underline{E_3(T)} = \frac{\{\boxed{\text{perimeter}} \text{ of } T \text{ near laser points (within 0.2 meters)}\}}{\{\text{total perimeter of } T\}}$$

$$E(T) = \underbrace{w_1E_1(T)}_{\text{Free space}} + \underbrace{w_2E_2(T)}_{\text{Laser points}} + \underbrace{w_3E_3(T)}_{\text{Regularization}}$$

3D CSG Reconstruction

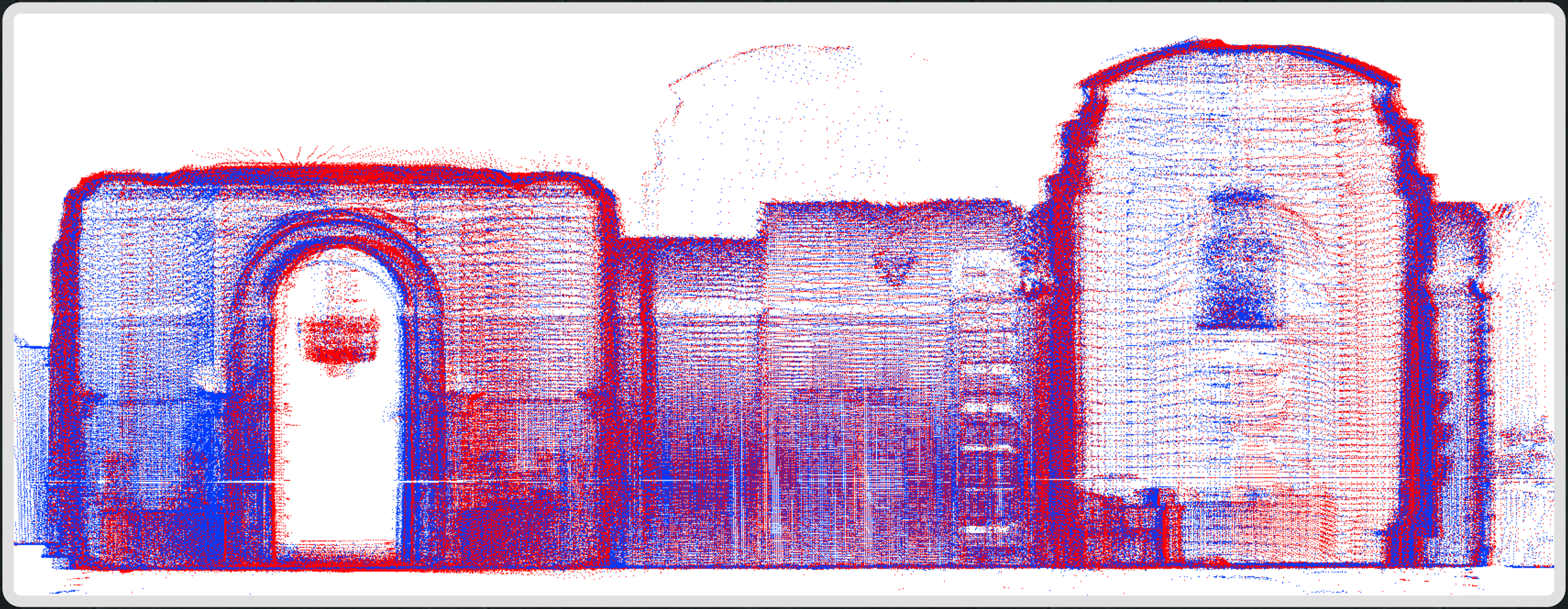
1. Generate primitives (cuboids)

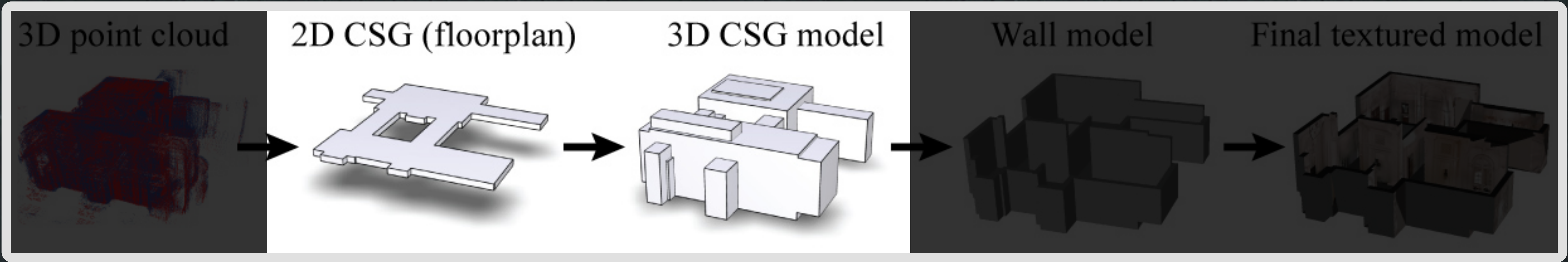
2. Choose a subset (out of primitive candidates)



3D CSG Reconstruction

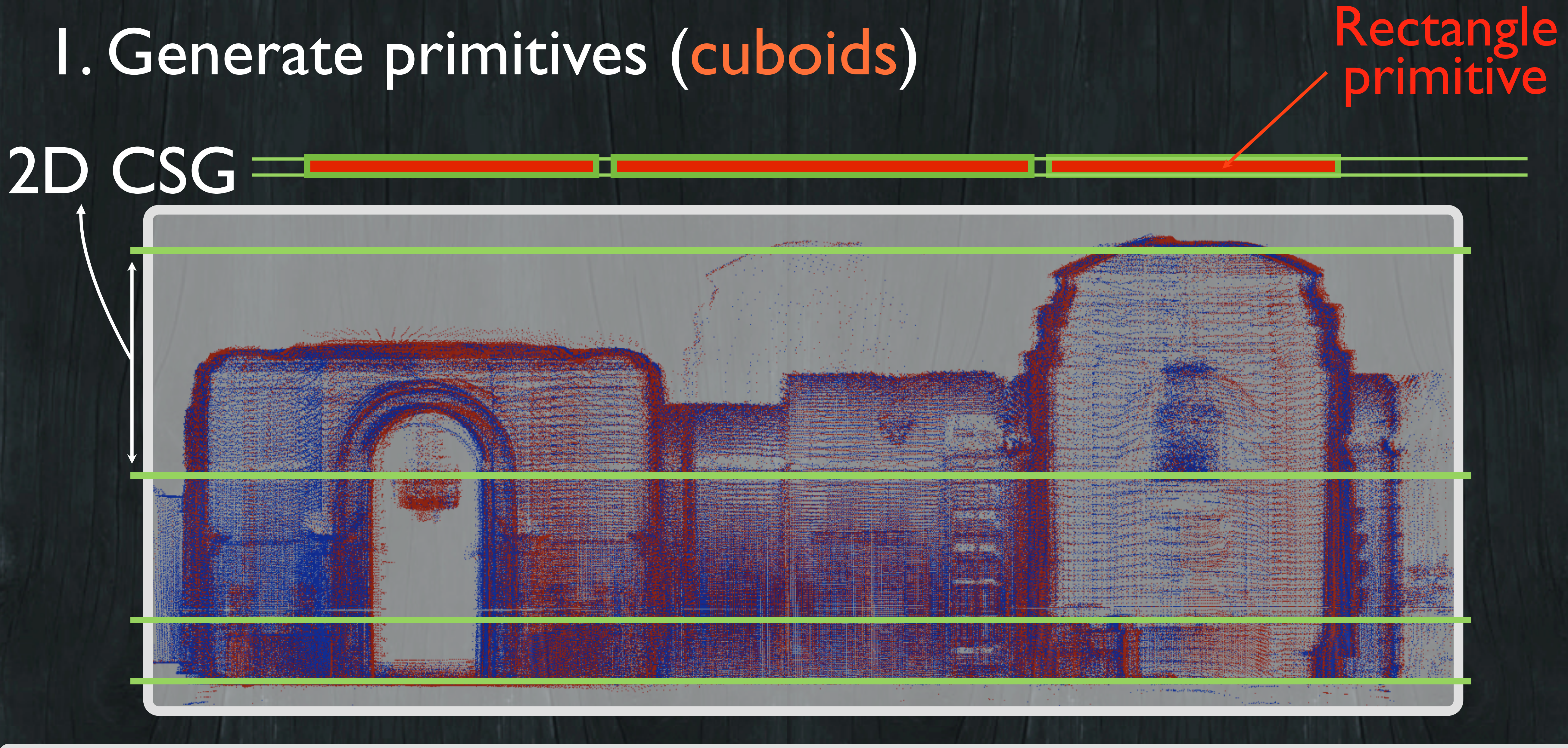
1. Generate primitives (cuboids)

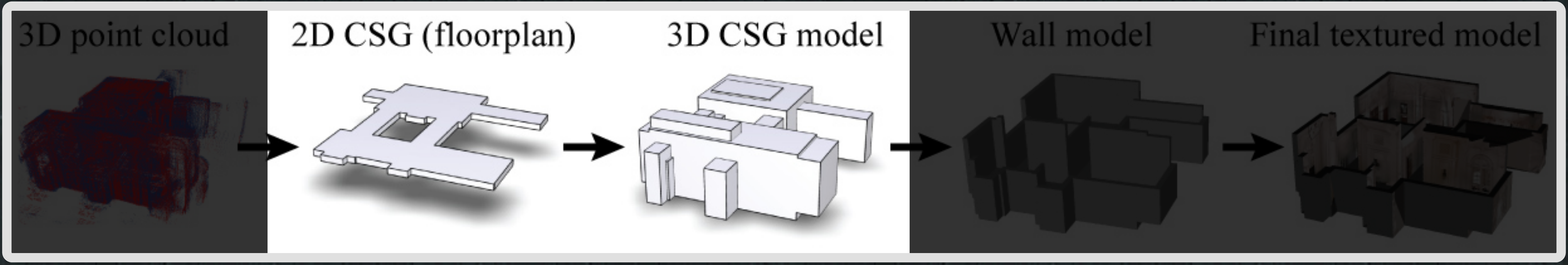




3D CSG Reconstruction

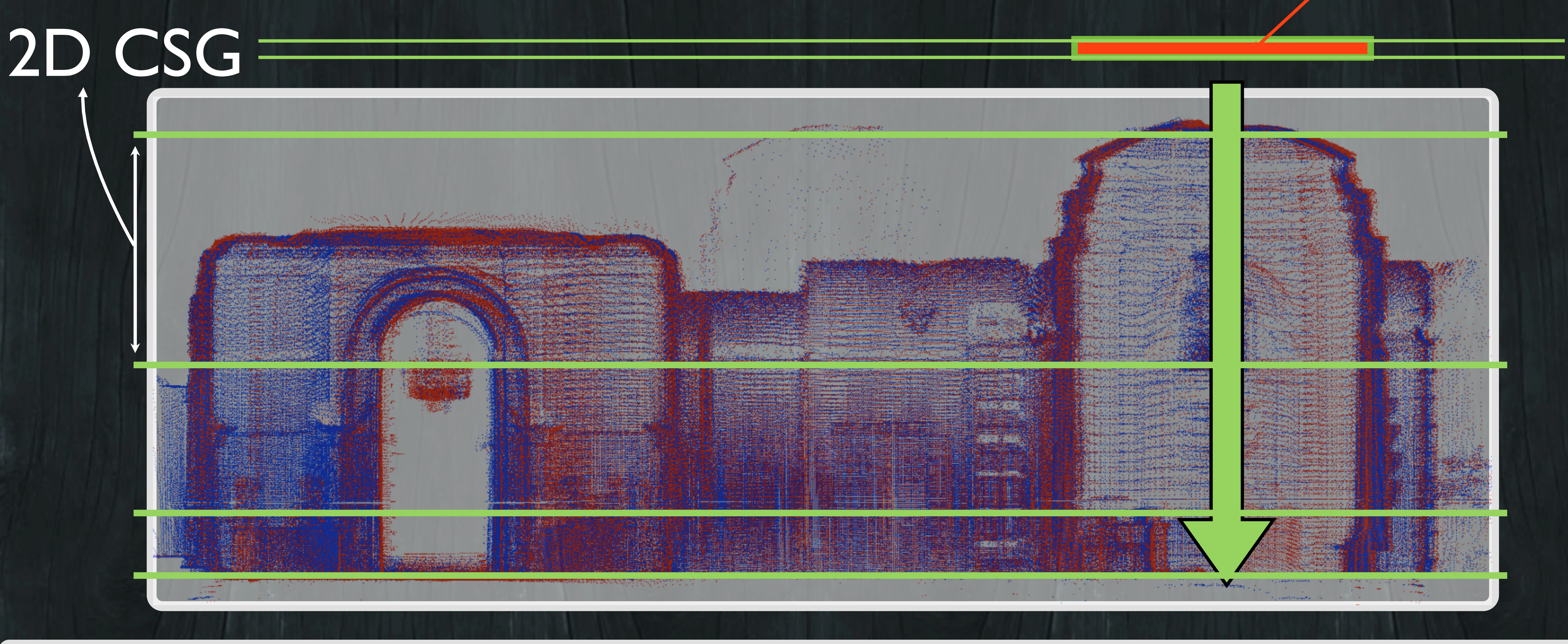
1. Generate primitives (cuboids)

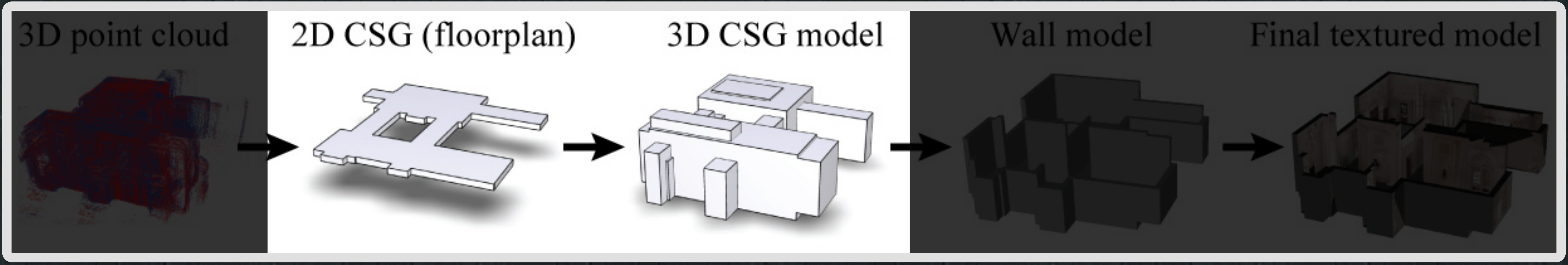




3D CSG Reconstruction

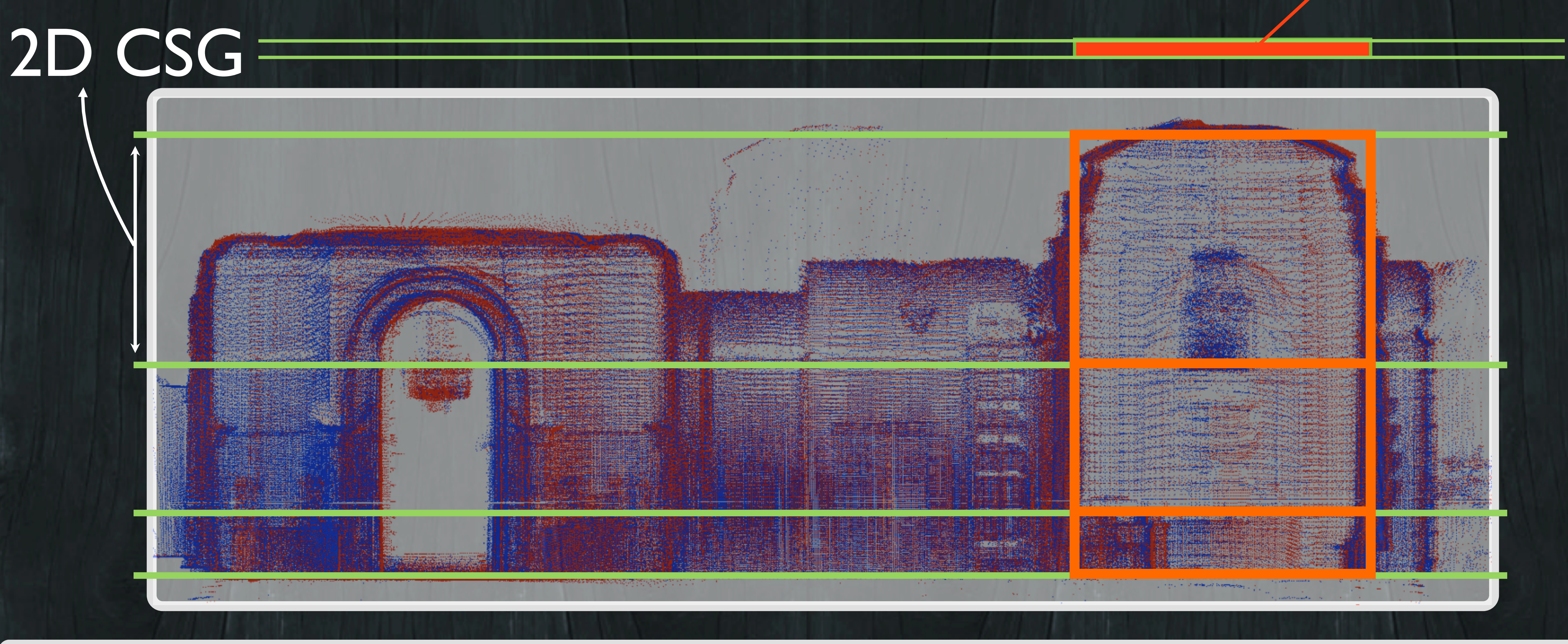
1. Generate primitives (cuboids)

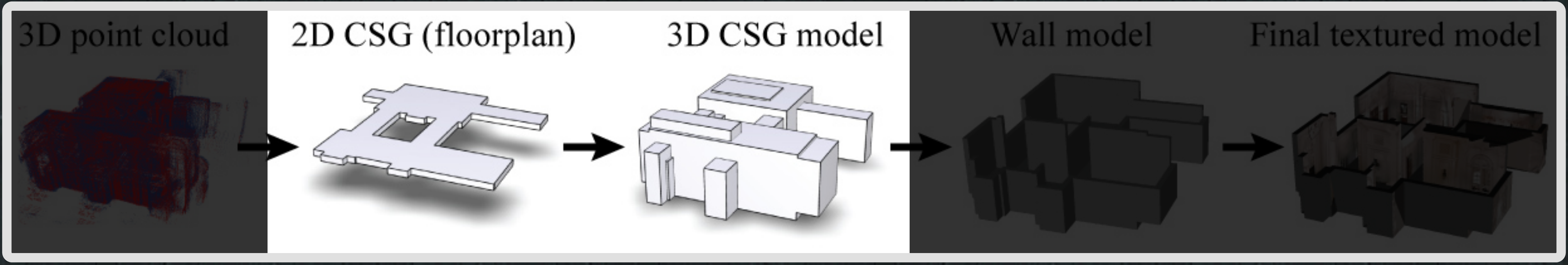




3D CSG Reconstruction

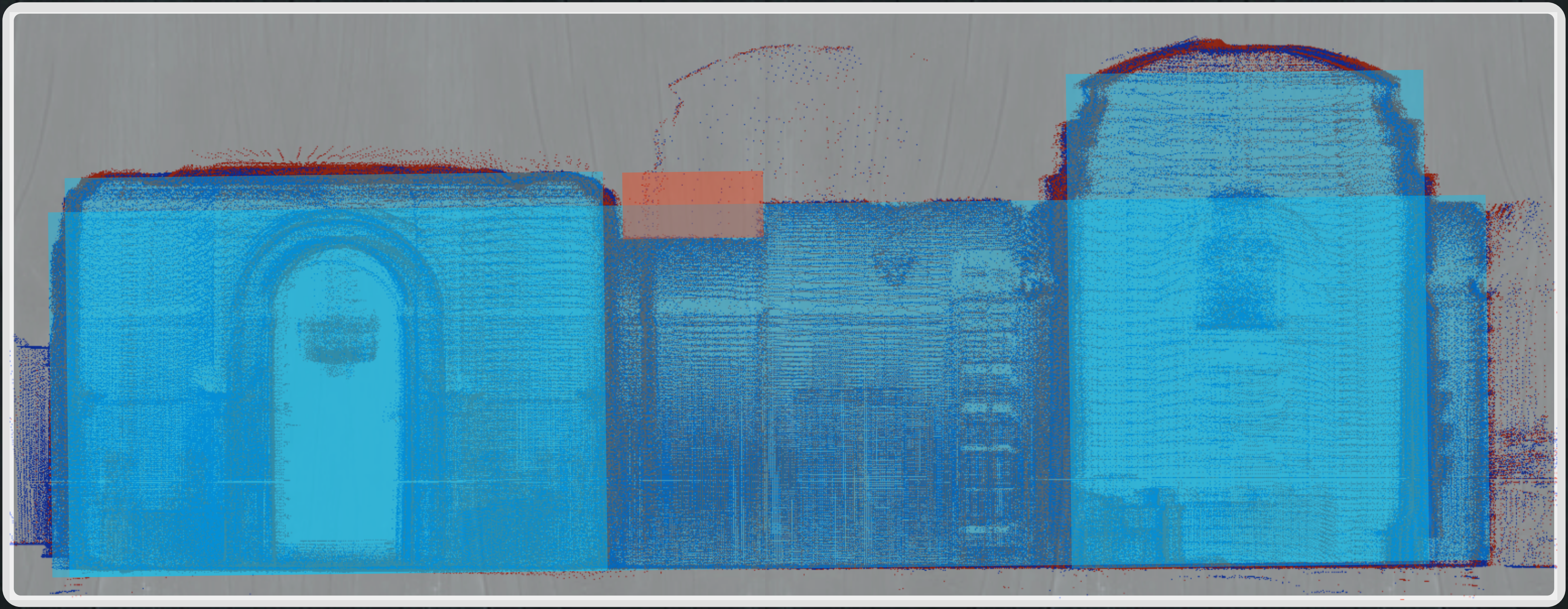
1. Generate primitives (cuboids)

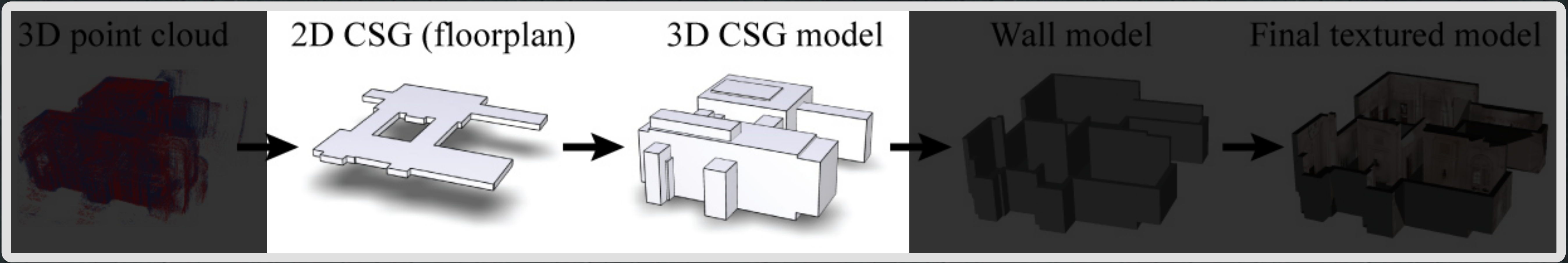




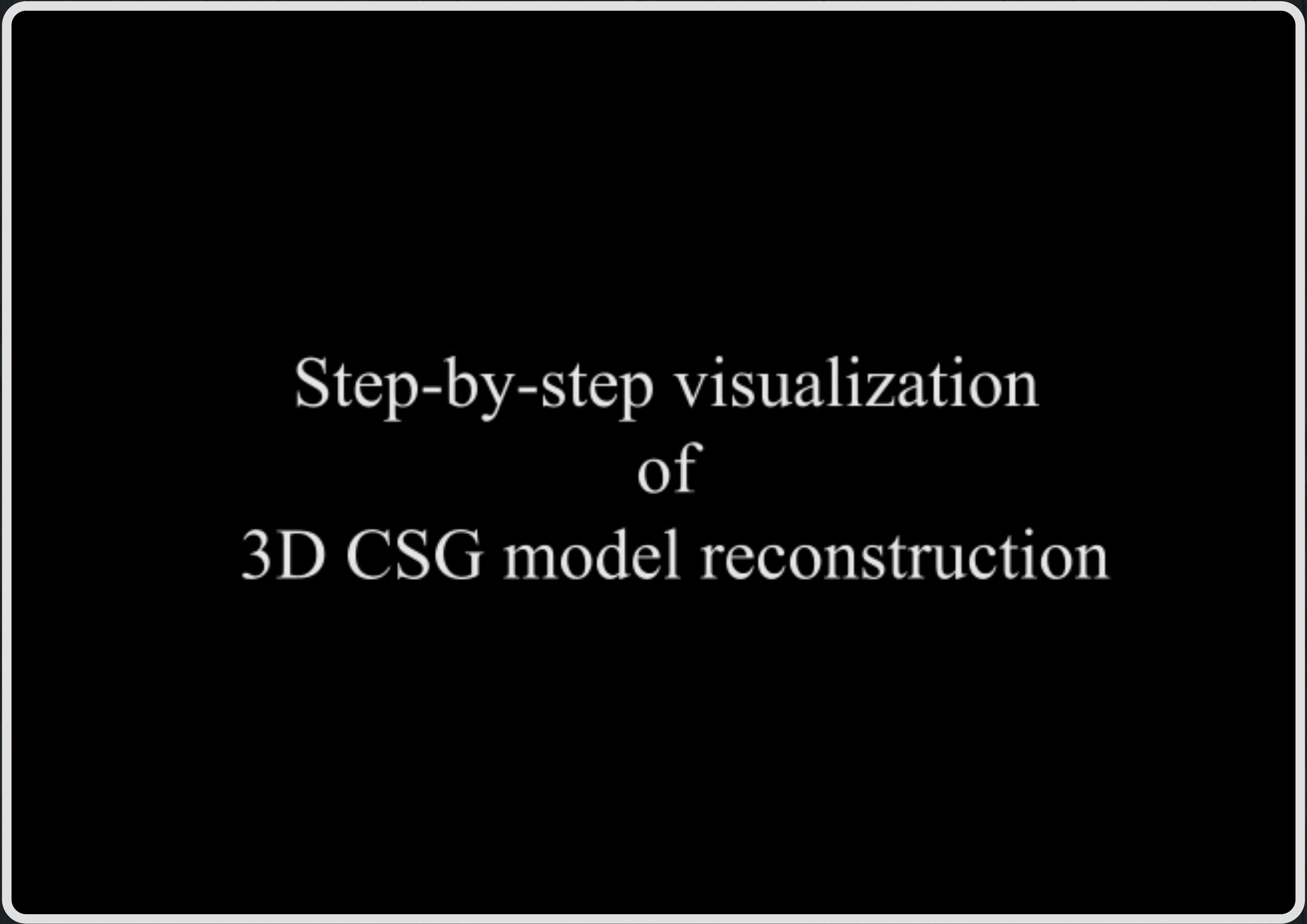
3D CSG Reconstruction

1. Generate primitives (cuboids)
2. Choose a subset



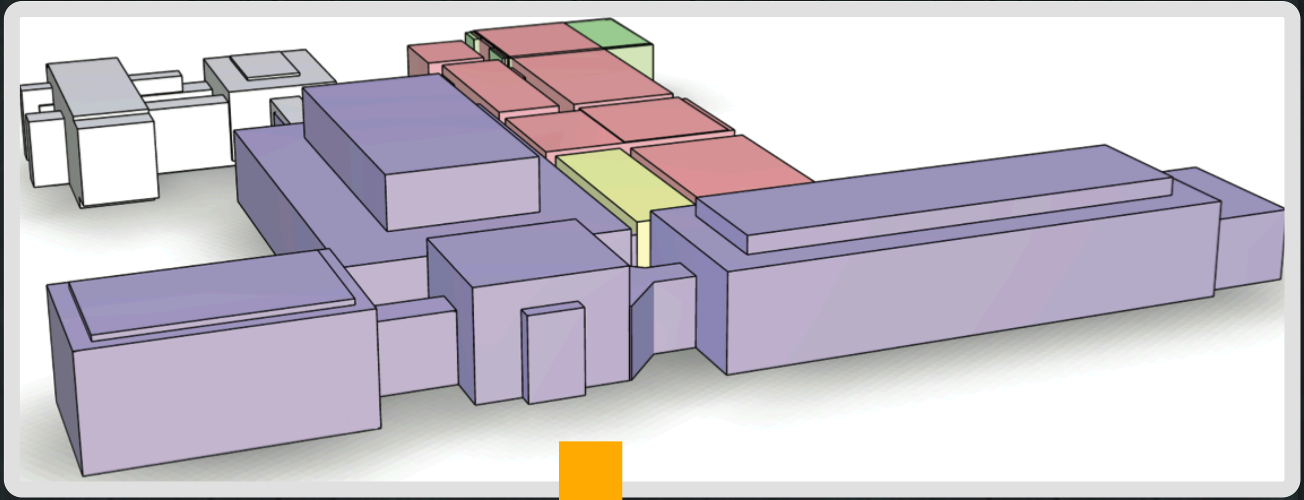


Algorithm on Run

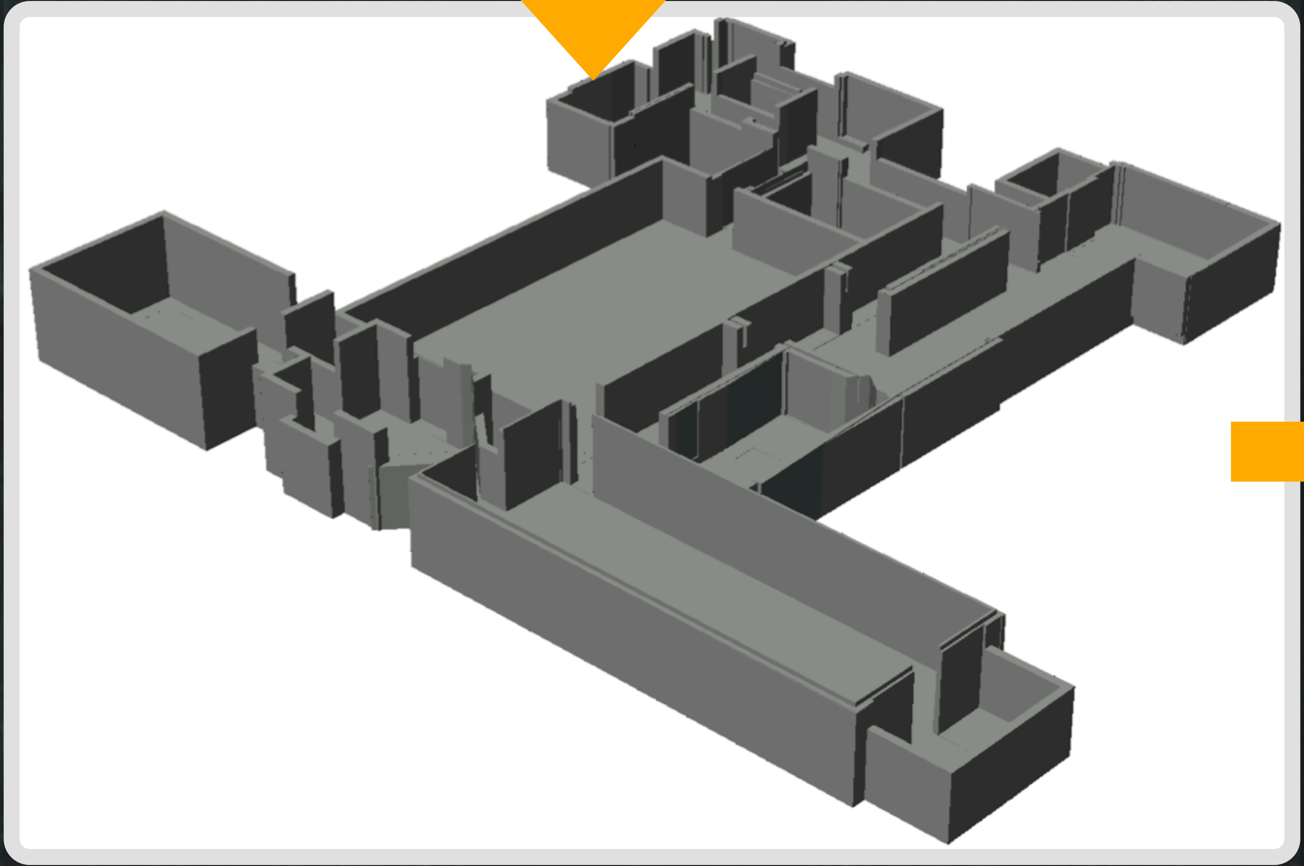


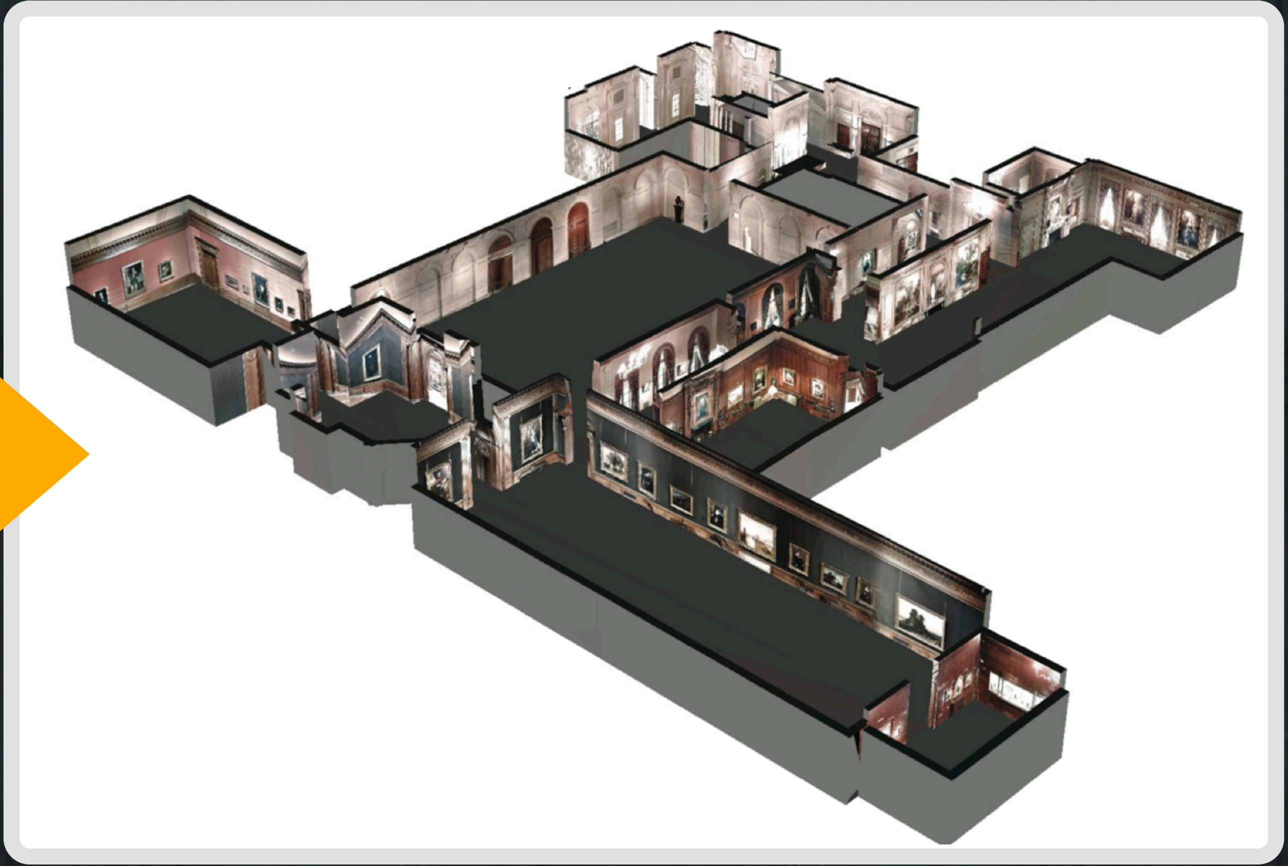
Step-by-step visualization
of
3D CSG model reconstruction

Last Step



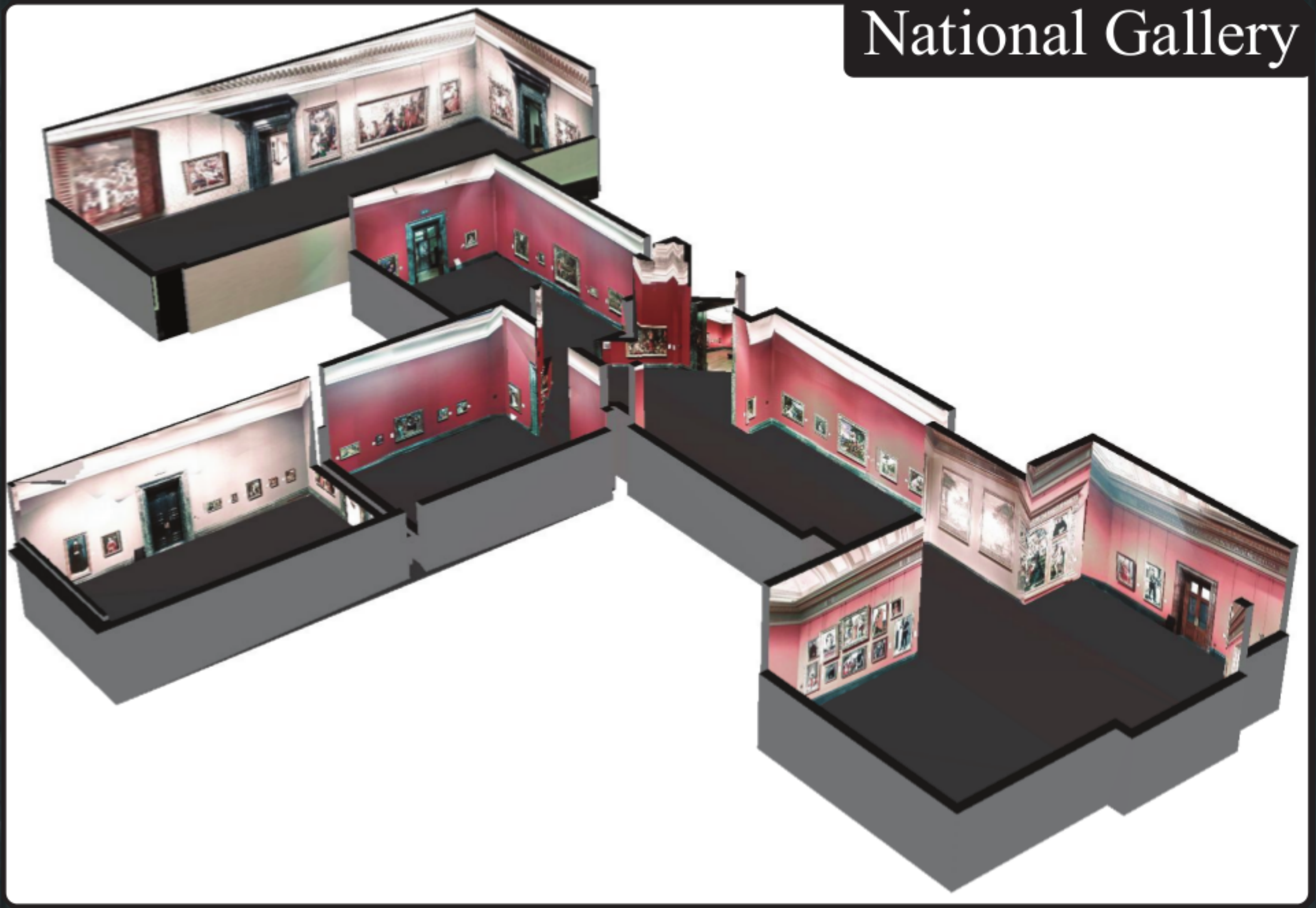
1. Remove Ceiling
2. Texture Mapping

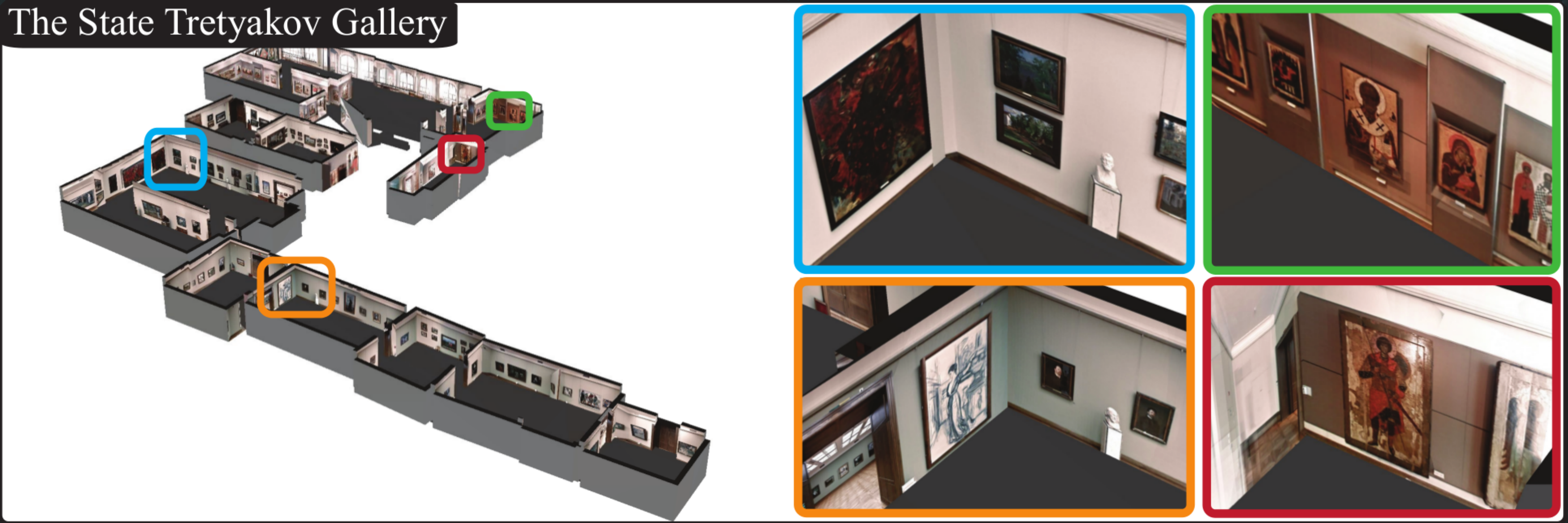




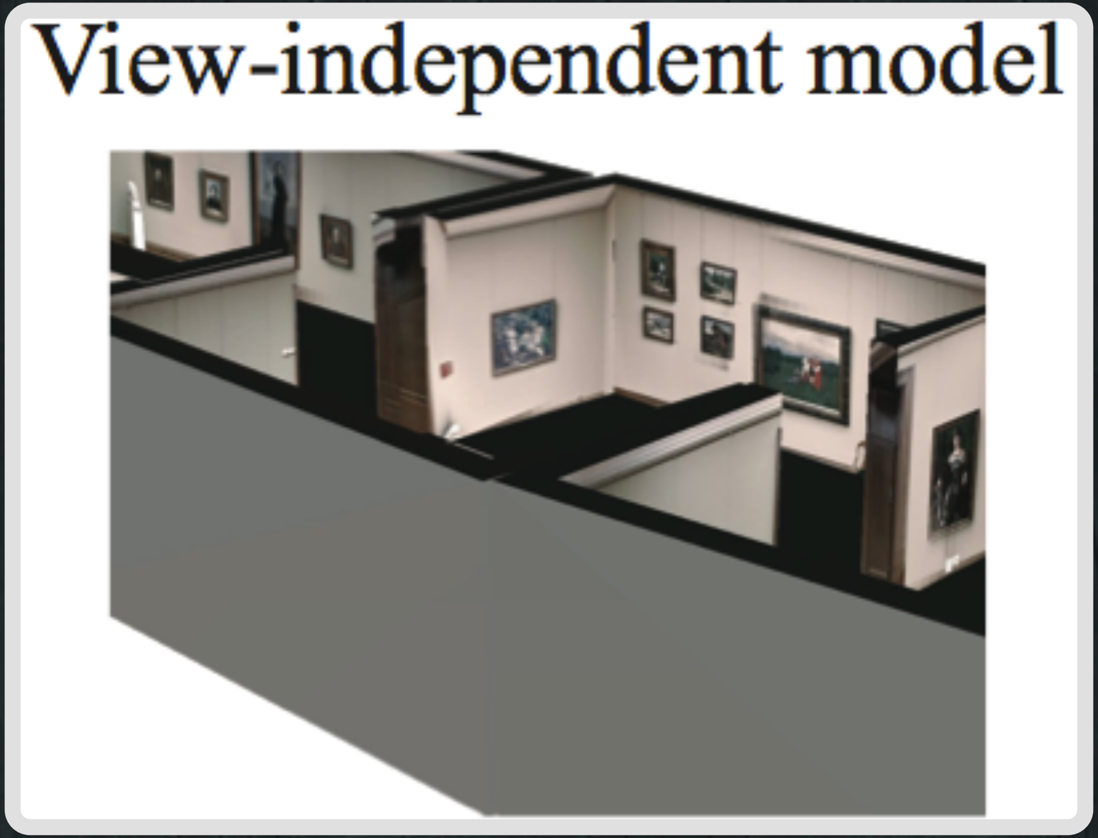
Per plane texture stitching [Xiao et al. 2008, Xiao et al. 2009].

The Frick Collection

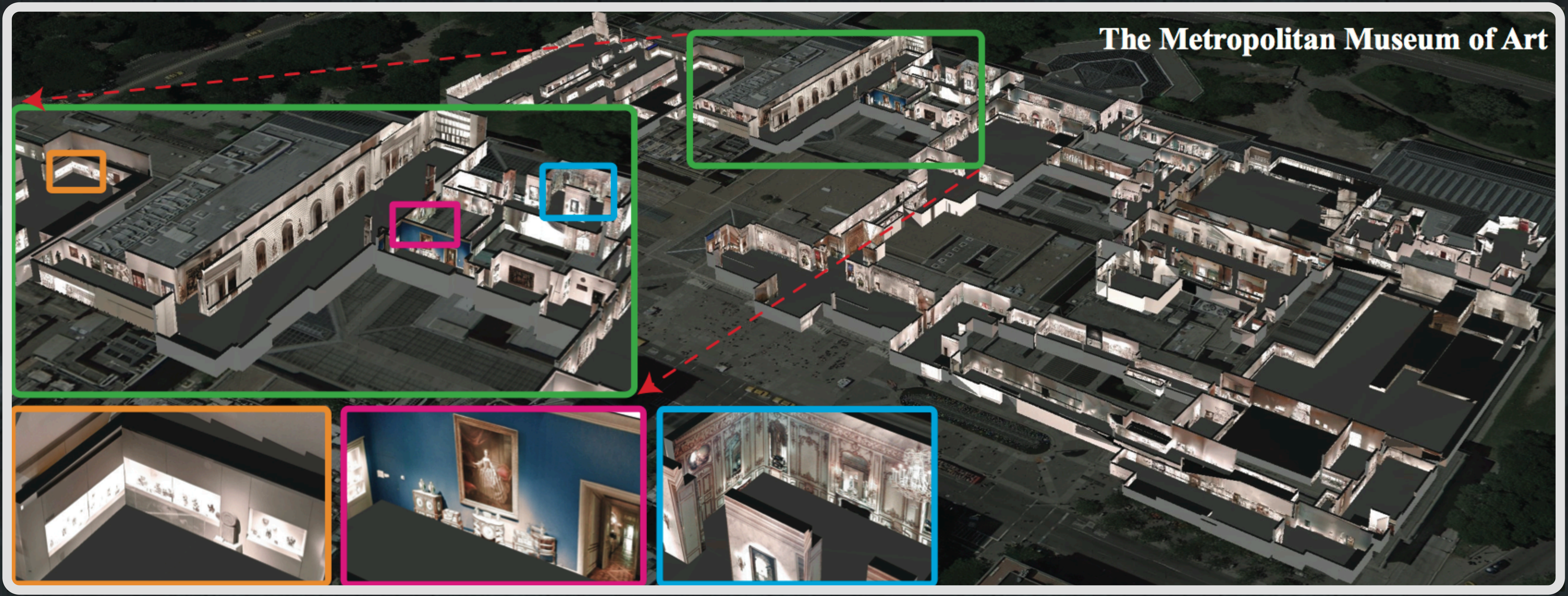

National Gallery


The State Tretyakov Gallery

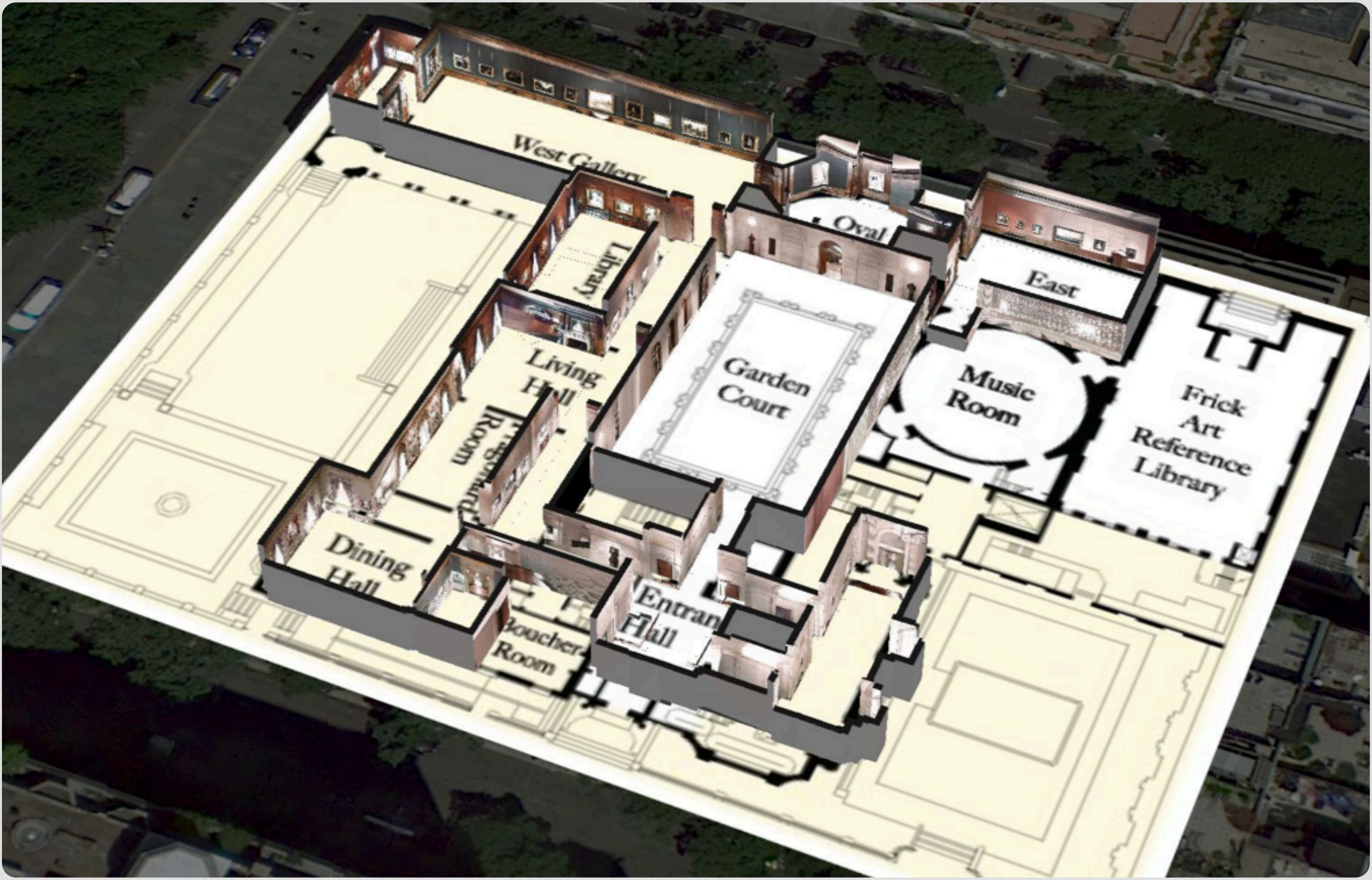
View-dependent Construction



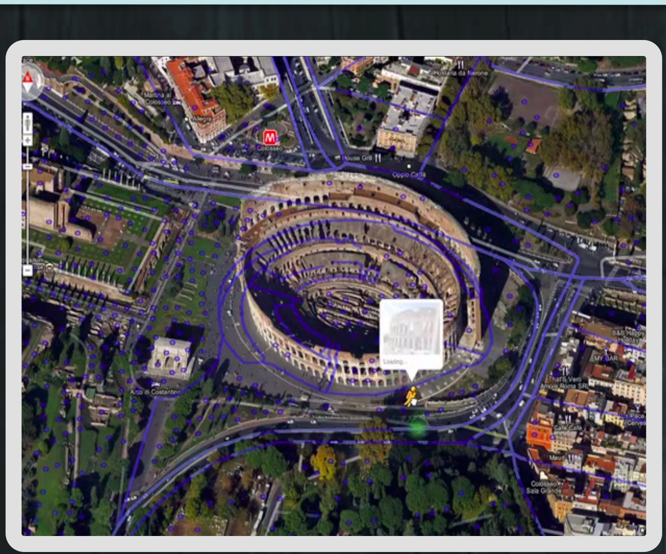
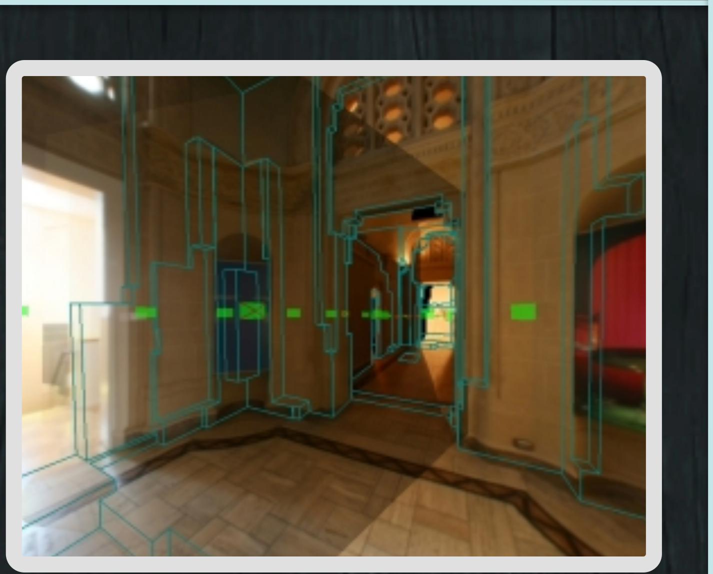
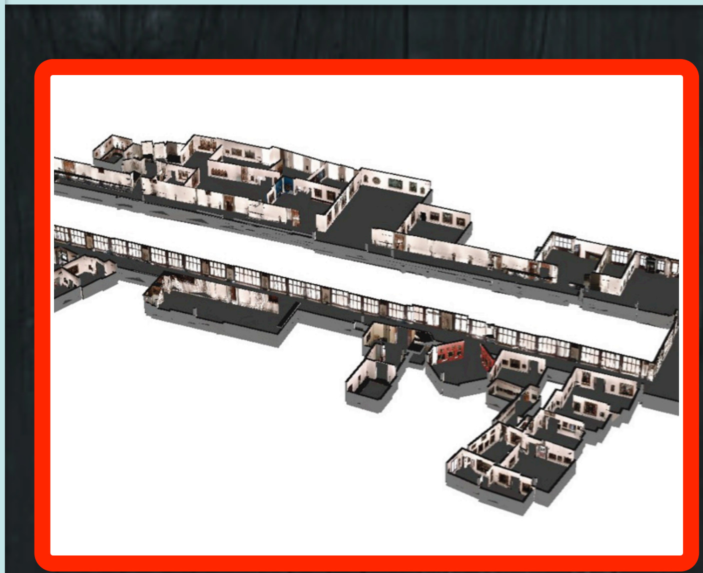
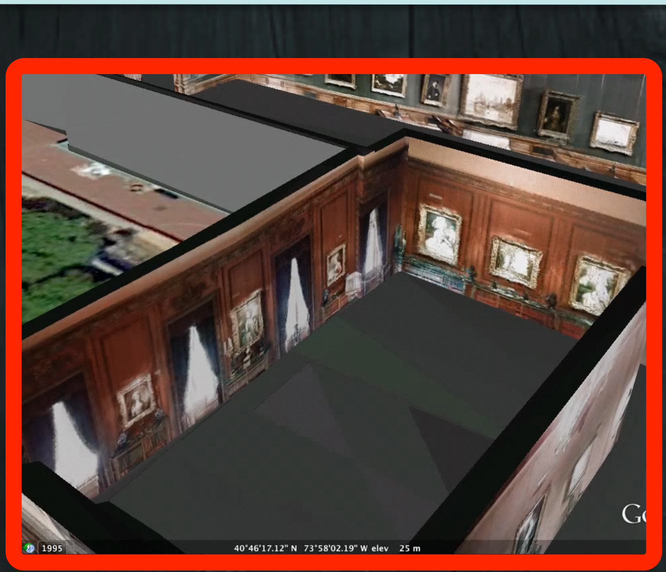
Aerial Maps



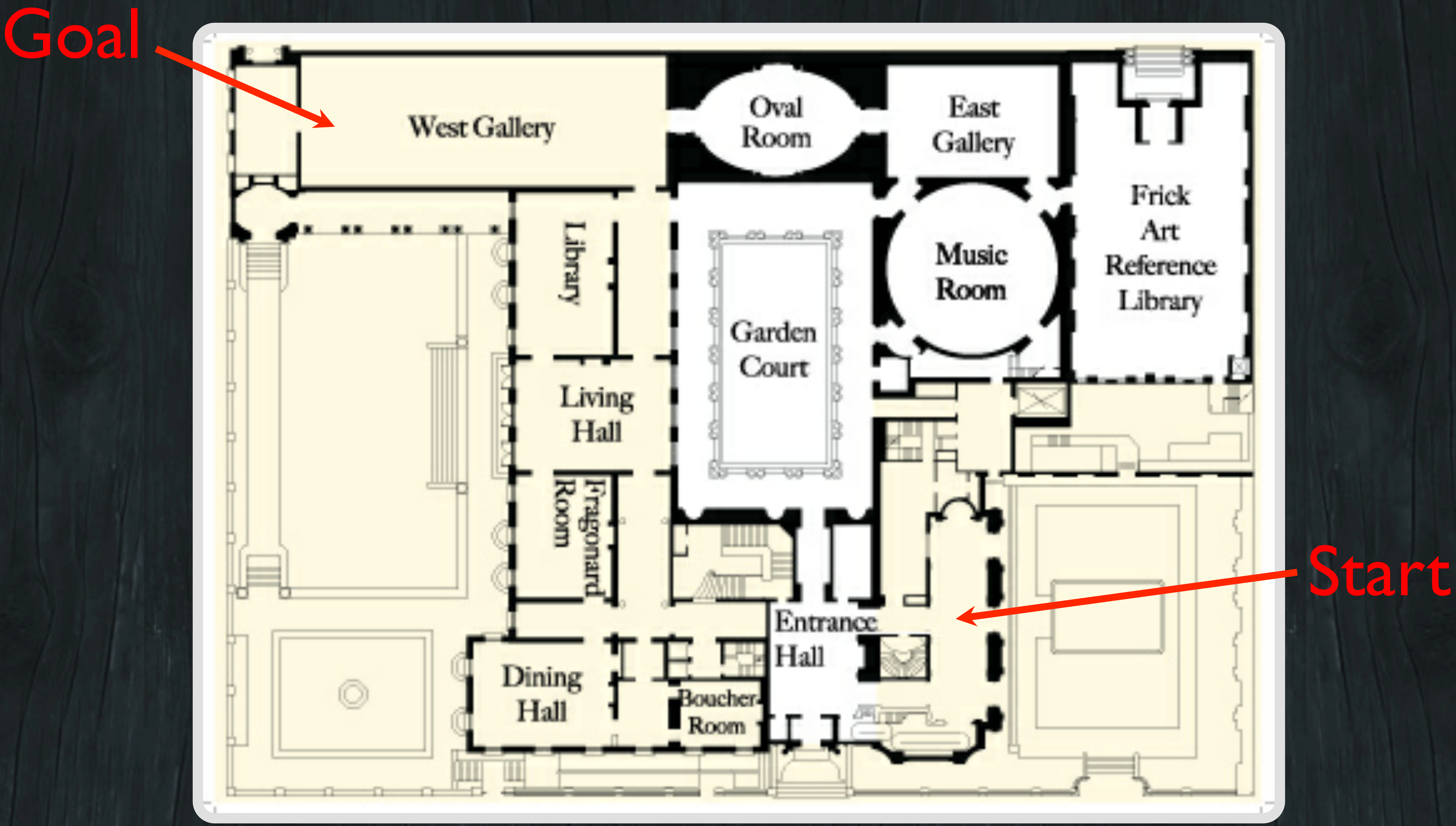
Hybrid Maps


West Gallery
Library
Oval
East
Living Hall
Garden Court
Music Room
Frick Art Reference Library
Fragonard Room
Dining Hall
Entrance Hall
Boucher Room

Ground vs. Aerial → Ground+Aerial

	Ground	Aerial	Ground+Aerial
Outdoor	 Google Streetview	 Google/Bing/NASA …	 Google MapsGL
Indoor	 Furukawa et al.	 This paper	 This paper

Enhance Navigation



Frick Collection Gallery (New York City)

5x

The Frick Collection
Exit Street View
COATROOM
© 2012 Google
Google earth

Take Home Message

Take Home Message

Technical Contribution

Conceptual Contribution

Take Home Message

Technical Contribution

Inverse CSG for
Large-scale Reconstruction

Conceptual Contribution

Take Home Message

Technical Contribution

Inverse CSG for
Large-scale Reconstruction

Conceptual Contribution

Indoor Photorealistic Maps
+ Aerial → Ground Transition
for Effective Navigation

Take Home Message

Technical Contribution

Inverse CSG for
Large-scale Reconstruction

Reconstruction

Conceptual Contribution

Indoor Photorealistic Maps
+ Aerial → Ground Transition
for Effective Navigation

Visualization

Take Home Message

Technical Contribution

Inverse CSG for
Large-scale Reconstruction

Reconstruction

Let machine see

Conceptual Contribution

Indoor Photorealistic Maps
+ Aerial → Ground Transition
for Effective Navigation

Visualization

Let human see better

Goals of Computer Vision

Take Home Message

Technical Contribution

Inverse CSG for
Large-scale Reconstruction

Reconstruction

Let machine see

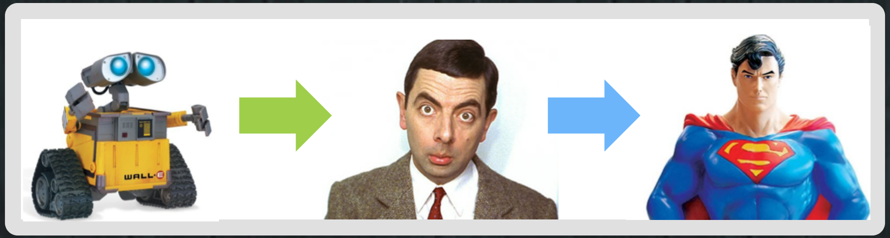
Conceptual Contribution

Indoor Photorealistic Maps
+ Aerial → Ground Transition
for Effective Navigation

Visualization

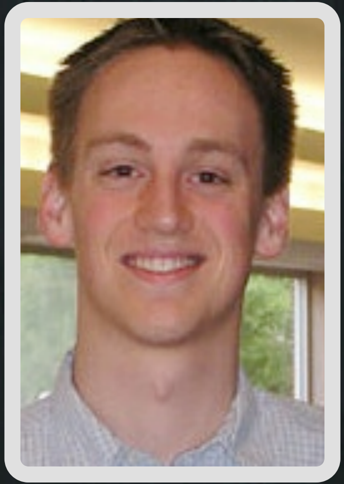
Let human see better

Goals of Computer Vision



Acknowledgements

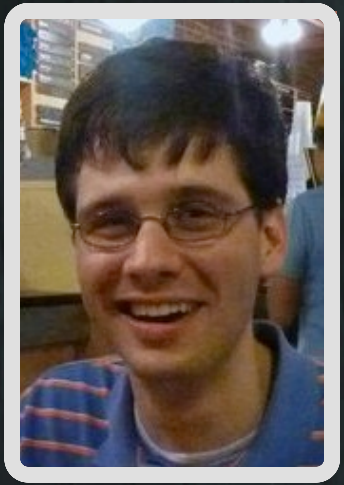
Steve Seitz



David Gallup

Carlos Hernandez

Maneesh Agrawala



Andrew Owens

Antonio Torralba

powered by Google Art Project

Research at Google

PhD Fellowship Program

ECCV 2012

Conference Travel Grant

Visart

Where Computer Vision Meets Art

Where is “the birth of venus”? Google Art from bird’s-eye view

Jianxiong Xiao

Time: Friday Oct 12, 2:30PM

Location: Room C 2F Affari

Demo: http://mit.edu/jxiao/museum/

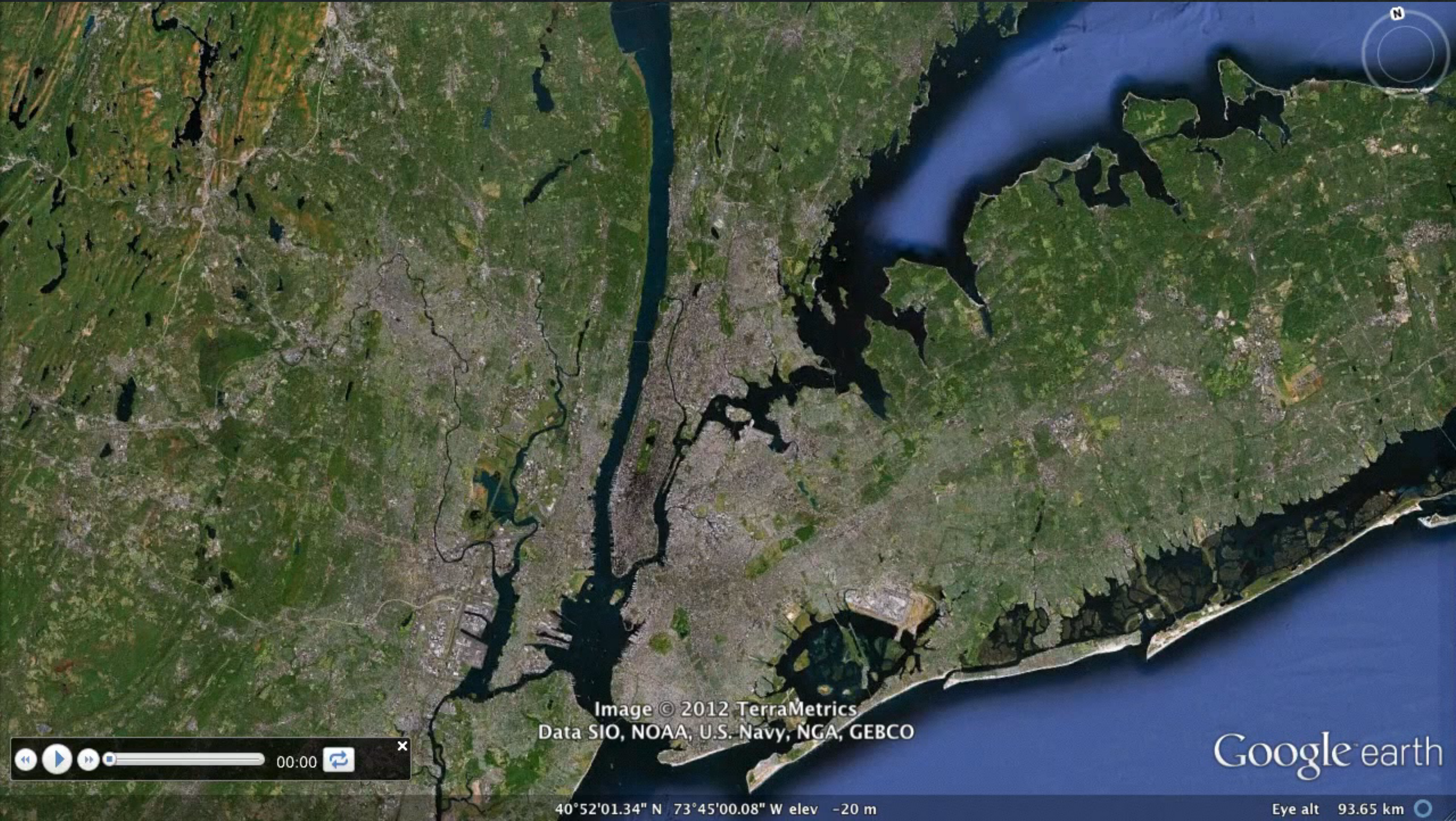


Image © 2012 TerraMetrics
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Google earth
00:00
40°52'01.34" N 73°45'00.08" W elev -20 m
Eye alt 93.65 km
N