# Visual Localization using Imperfect 3D Models from the Internet

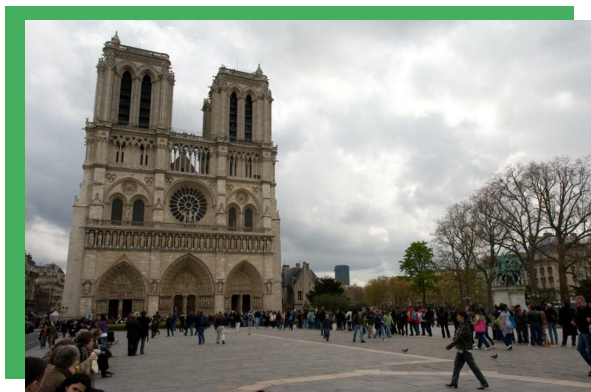**Vojtech Panek[1,2]**　　**Zuzana Kukelova[3]**　　**Torsten Sattler[2]**

[1]Faculty of Electrical Engineering, Czech Technical University (CTU) in Prague
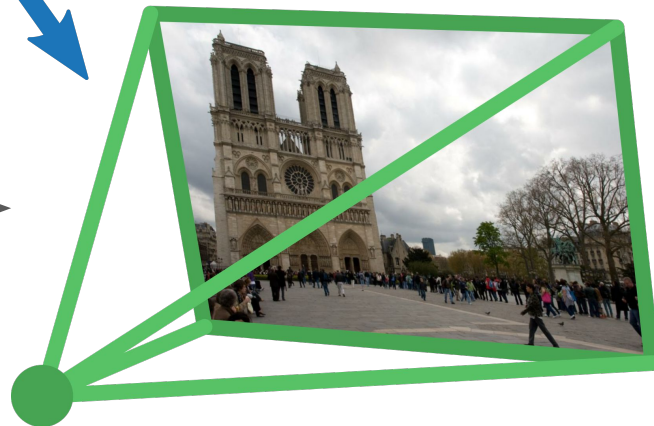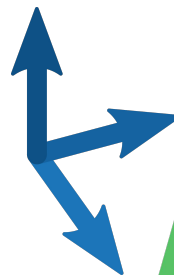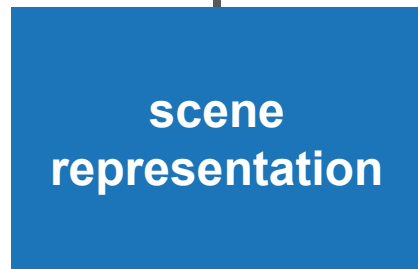[2]Czech Institute of Informatics, Robotics and Cybernetics, CTU in Prague
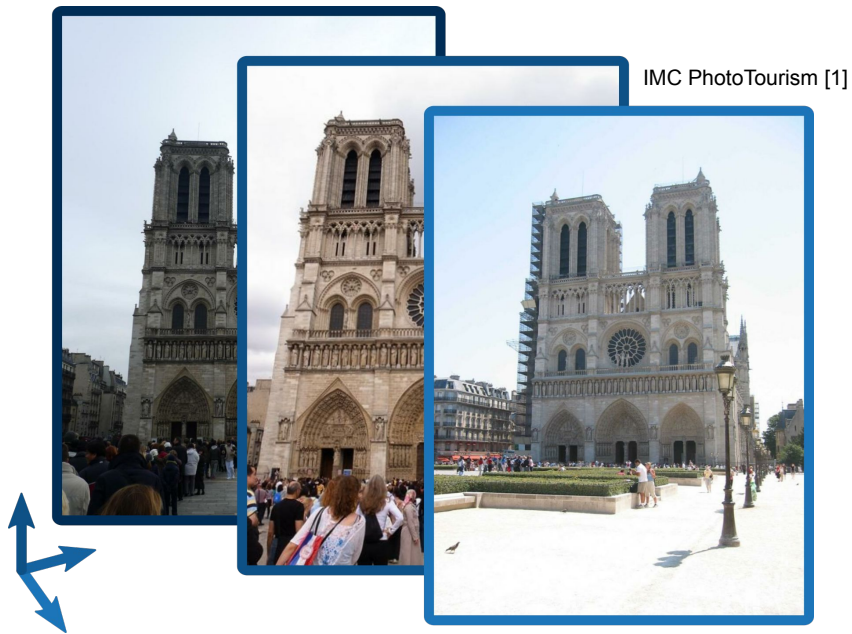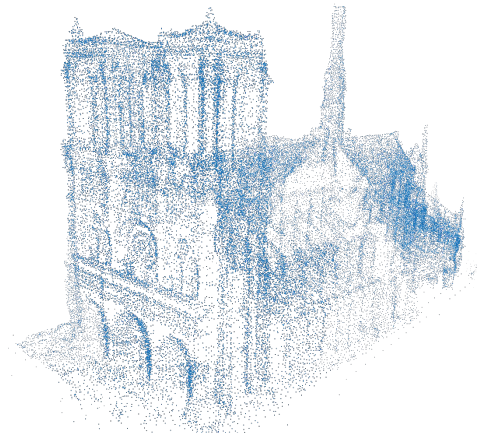[3]Visual Recognition Group, Faculty of Electrical Engineering, CTU in Prague

**WED-PM-078**

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

IMC PhotoTourism [1]

**query image**

**localization pipeline**

**scene representation**

**estimated camera pose**

[1] Jin, Y., Mishkin, D., Mishchuk, A., Matas, J., Fua, P., Yi, K. M., Trulls, E. "Image matching across wide baselines: From paper to practice." IJCV, 2021.

IMC PhotoTourism [1]

**SfM point clouds**

**reference images**

Miguel Bandera
(https://skfb.ly/6QWu7)

**MVS meshes**

IMC PhotoTourism [1]
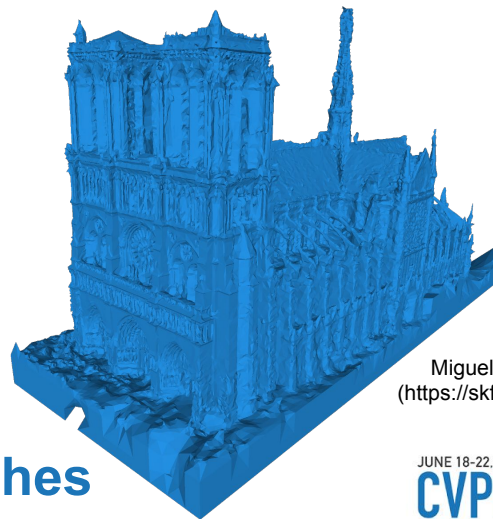
MiniWorld3D
(https://mmf.io/o/91899)
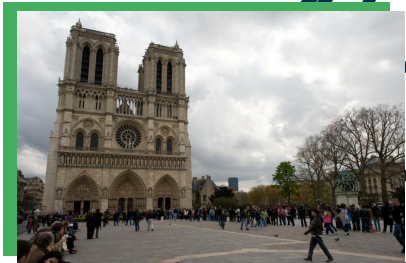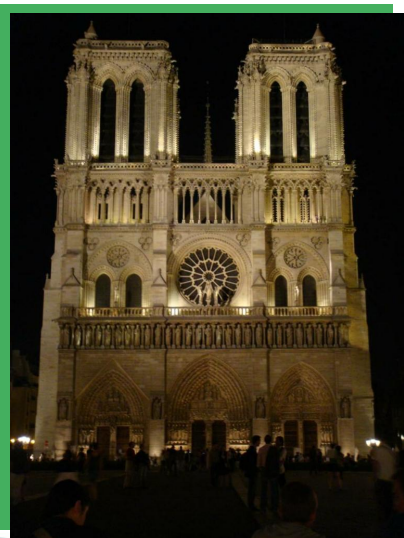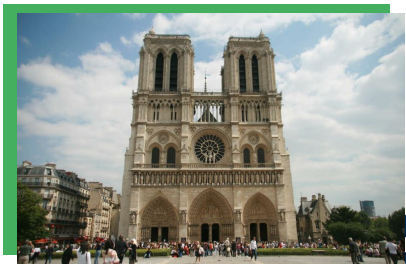
**query images**

model sharing website
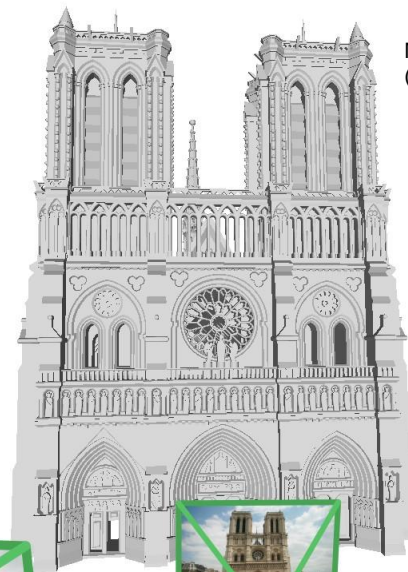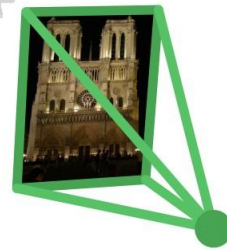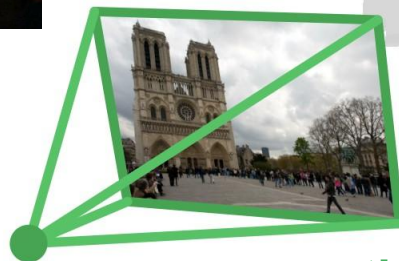
**estimated camera poses**

**Create AR experiences without data acquisition**

CZECH INSTITUTE OF INFORMATICS ROBOTICS AND CYBERNETICS CTU IN PRAGUE

FACULTY OF ELECTRICAL ENGINEERING CTU IN PRAGUE

JUNE 18-22, 2023
CVPR VANCOUVER, CANADA
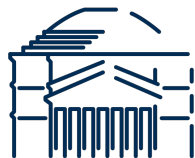
Notre Dame     Pantheon     Reichstag     St. Peter's Square     St. Vitus Cathedral

Miguel Bandera
(https://skfb.ly/6QWu7)
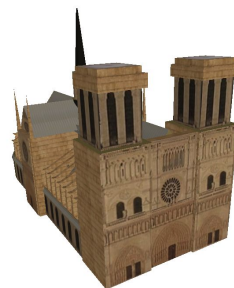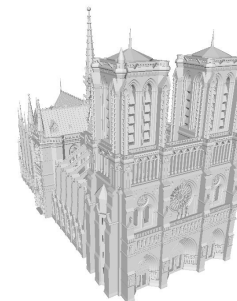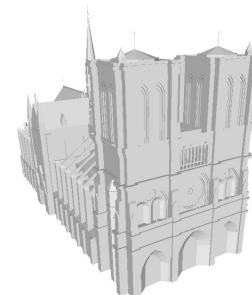
Chigirinsky
(https://skfb.ly/6Rn9M)

Alejandro Diaz
(https://skfb.ly/31dba)

Little-Goomba
(https://bit.ly/3QWeOxY)

MiniWorld3D
(https://mmf.io/o/91899)

giotis
(https://bit.ly/3QOTQ41)

reference
MVS mesh

Fovea
https://skfb.ly/6RZHt

brnipon
https://bit.ly/3CCwTwP

Ultima Ratio
https://bit.ly/3AuN2BK

Adsman007
https://bit.ly/3ASvl0b

Emanuele Viani
https://skfb.ly/EAKB
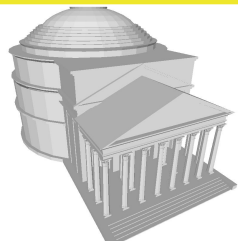
MiniWorld3D
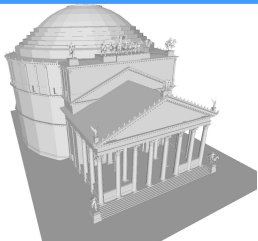(https://mmf.io/o/91899)
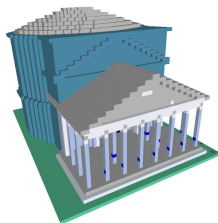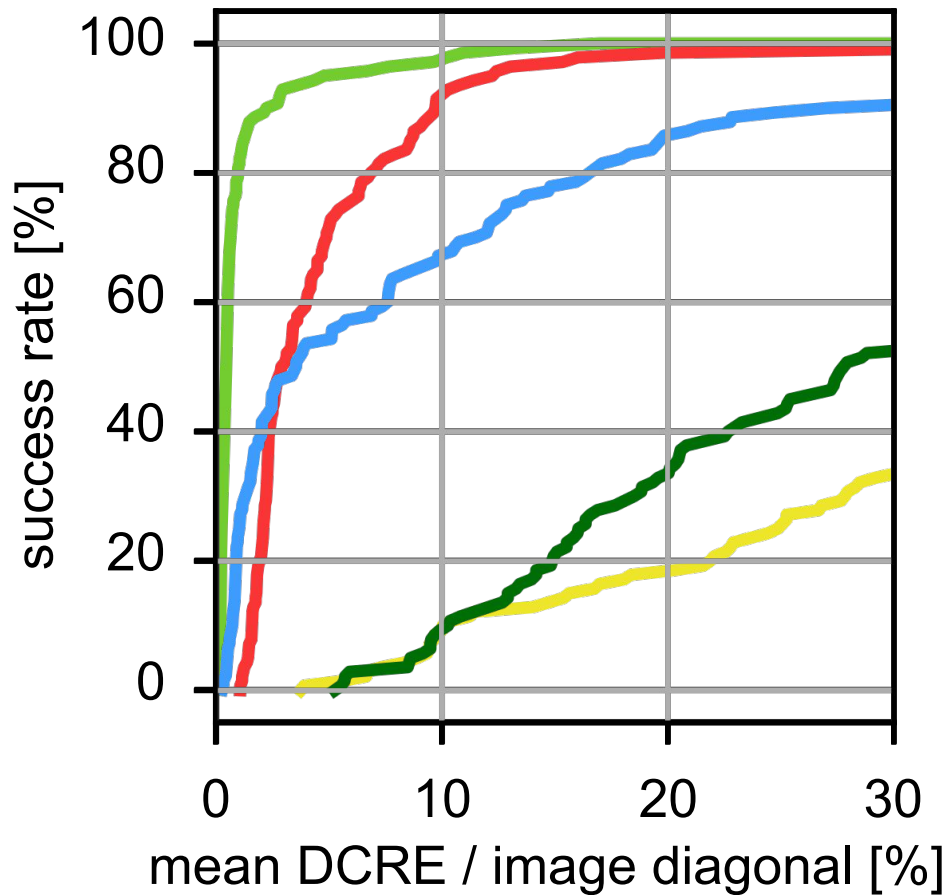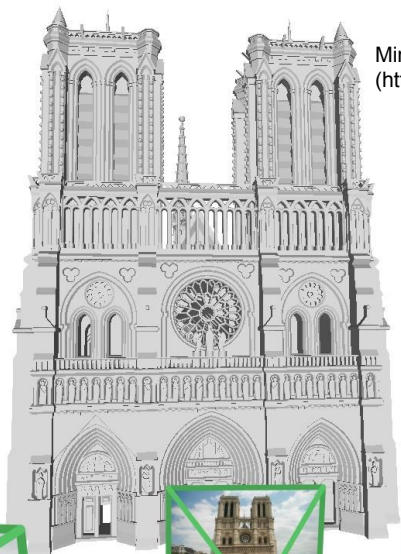
**estimated camera poses**

**MeshLoc [2]:** Localization on meshes can achieve SoTA results

- compact scene representations
- usable also for other tasks

[2] Panek, V., Kukelova, Z., & Sattler, T., "MeshLoc: Mesh-Based Visual Localization." ECCV, 2022.

MiniWorld3D
(https://mmf.io/o/91899)

**estimated camera poses**

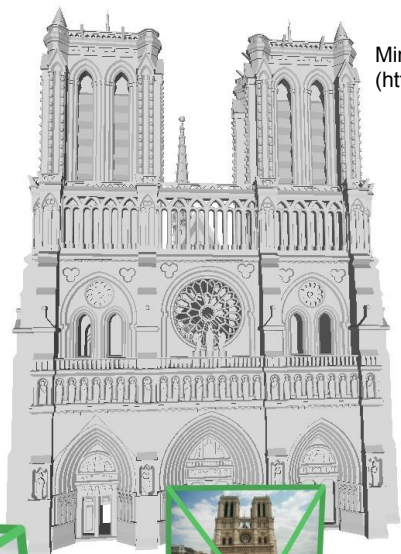**MeshLoc [2]:** Localization on meshes can achieve SoTA results
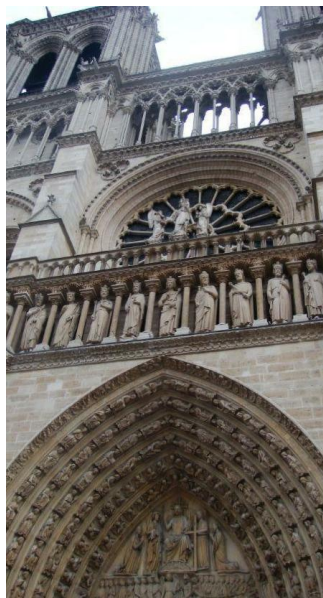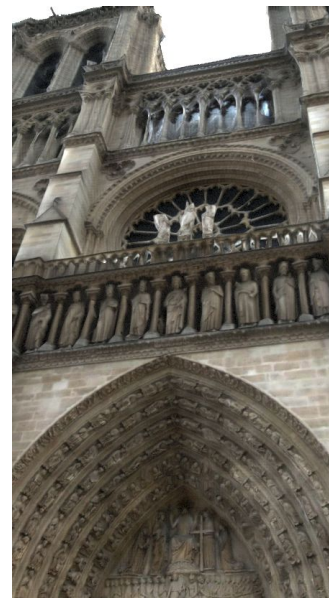
**photo**          **textured MVS**

Miguel Bandera (https://skfb.ly/6QWu7)
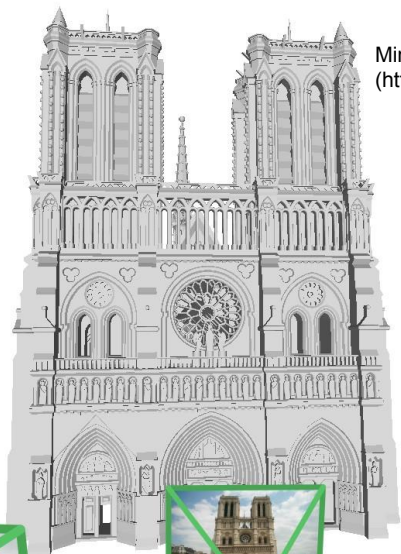
[2] Panek, V., Kukelova, Z., & Sattler, T., "MeshLoc: Mesh-Based Visual Localization." ECCV, 2022.

CZECH INSTITUTE OF INFORMATICS ROBOTICS AND CYBERNETICS CTU IN PRAGUE

FACULTY OF ELECTRICAL ENGINEERING CTU IN PRAGUE

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

MiniWorld3D
(https://mmf.io/o/91899)

**estimated camera poses**

**MeshLoc [2]:** Localization on meshes can achieve SoTA results
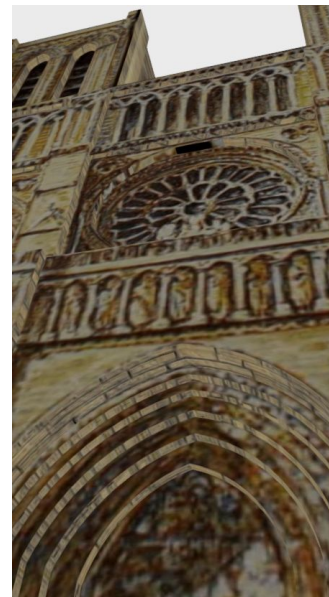
**photo**　　　　**hand-drawn CAD**

Alejandro Diaz (https://skfb.ly/31dba)

[2] Panek, V., Kukelova, Z., & Sattler, T., "MeshLoc: Mesh-Based Visual Localization." ECCV, 2022.

CZECH INSTITUTE
OF INFORMATICS
ROBOTICS AND
CYBERNETICS
CTU IN PRAGUE

FACULTY
OF ELECTRICAL
ENGINEERING
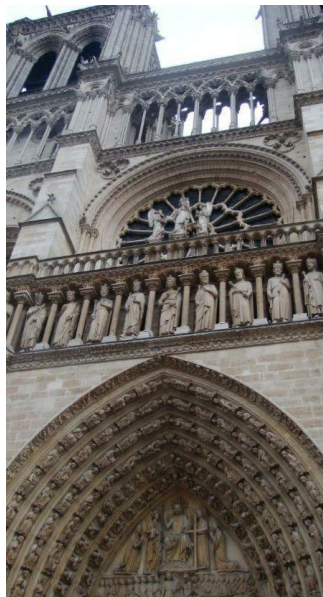CTU IN PRAGUE

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

MiniWorld3D
(https://mmf.io/o/91899)
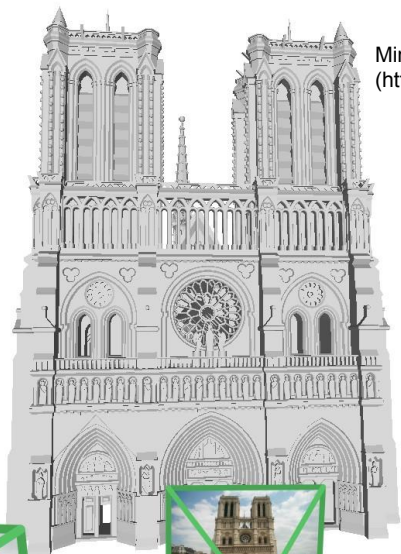
**estimated camera poses**

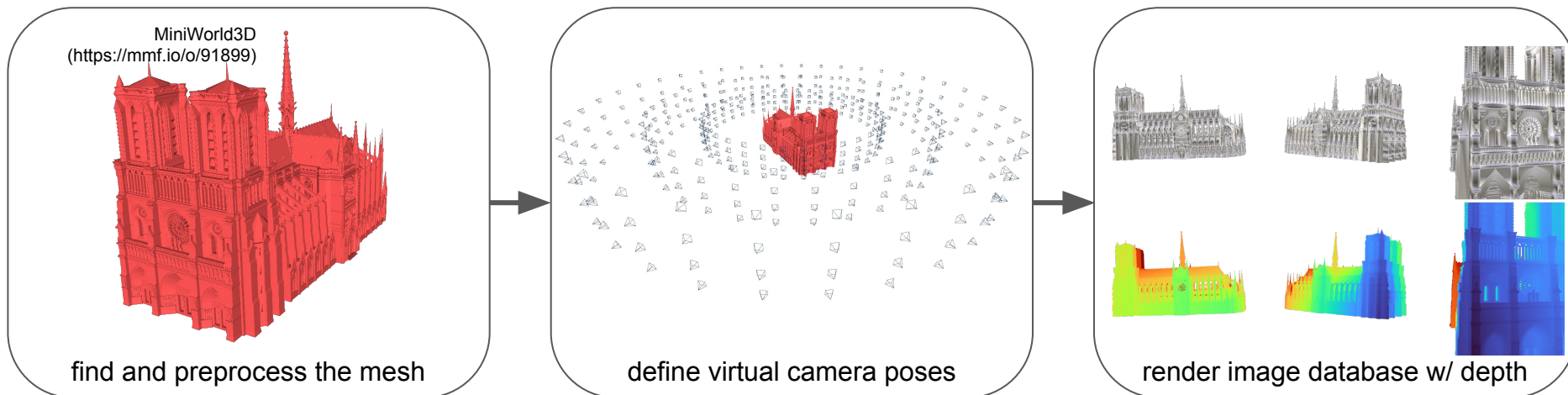**MeshLoc [2]:** Localization on meshes can achieve SoTA results

**Questions:**

Can you just download a 3D model from the Internet and localize with it?

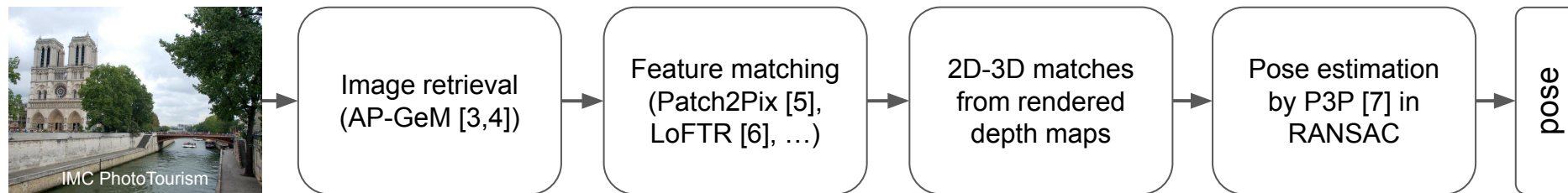How does the quality of the mesh influence the localization accuracy?

[2] Panek, V., Kukelova, Z., & Sattler, T., "MeshLoc: Mesh-Based Visual Localization." ECCV, 2022.

# Data preparation (offline)



MiniWorld3D
(https://mmf.io/o/91899)

find and preprocess the mesh

define virtual camera poses

render image database w/ depth

# Localization (online)
Use e.g. **MeshLoc [2]** (or other suitable pipeline) to localize



IMC PhotoTourism

Image retrieval
(AP-GeM [3,4])

Feature matching
(Patch2Pix [5],
LoFTR [6], …)

2D-3D matches
from rendered
depth maps

Pose estimation
by P3P [7] in
RANSAC

pose

[3] Gordo, A., Almazán, J., Revaud, J., Larlus, D. "End-to-End Learning of Deep Visual Representations for Image Retrieval." IJCV, 2016.
[4] Revaud, J., Almazán, J., Rezende, R.S., & Souza, C.R. "Learning With Average Precision: Training Image Retrieval With a Listwise Loss." ICCV, 2019.
[5] Zhou, Q., Sattler, T., & Leal-Taixé, L., "Patch2Pix: Epipolar-Guided Pixel-Level Correspondences." CVPR, 2021.
[6] Sun, J., Shen, Z., Wang, Y., Bao, H., & Zhou, X., "LoFTR: Detector-Free Local Feature Matching with Transformers." CVPR, 2021.
[7] Persson, M., Nordberg, K. "Lambda Twist: An Accurate Fast Robust Perspective Three Point (P3P) Solver." ECCV, 2018.

JUNE 18-22, 2023
CVPR VANCOUVER, CANADA

# Evaluation - ground truth pose definition
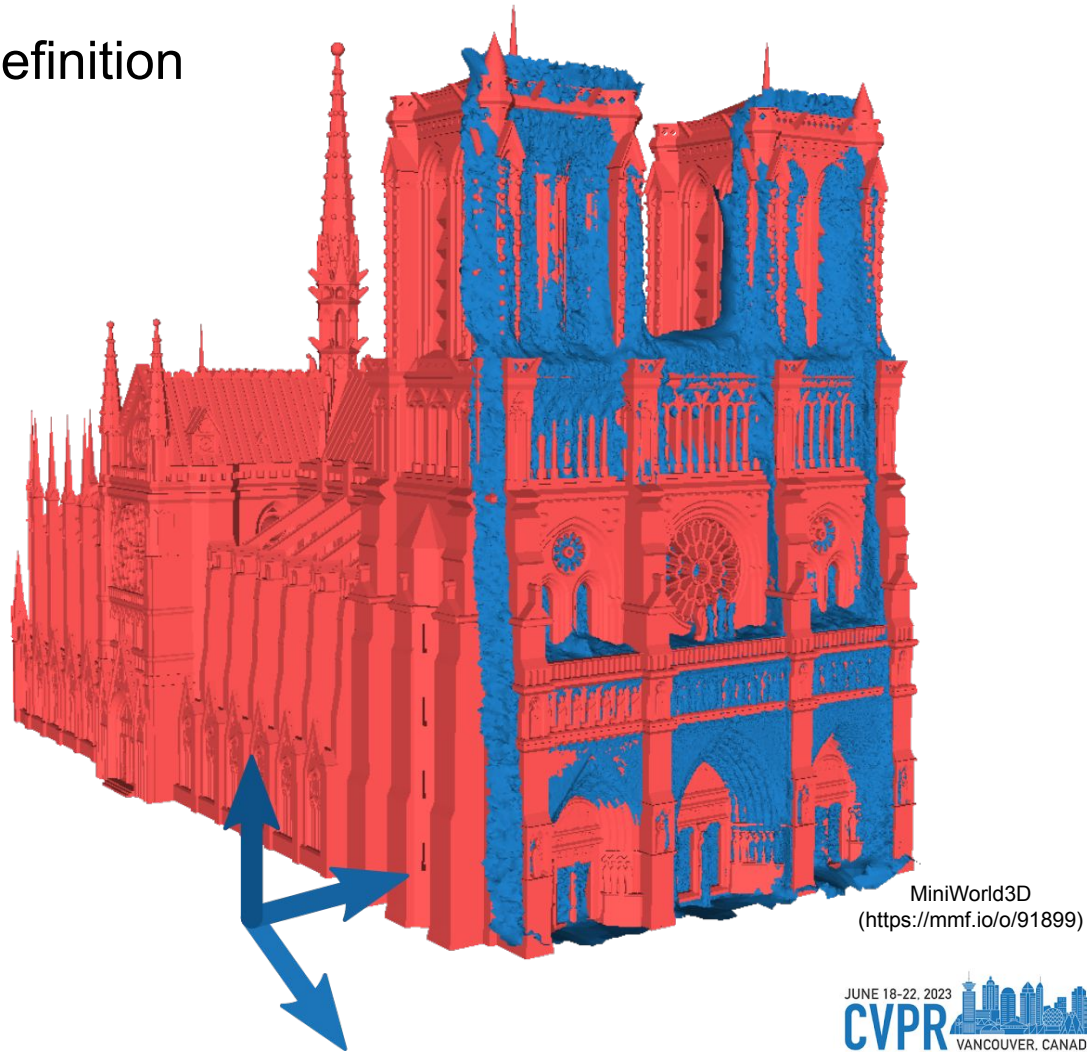


IMC PhotoTourism [1]

**reference images**
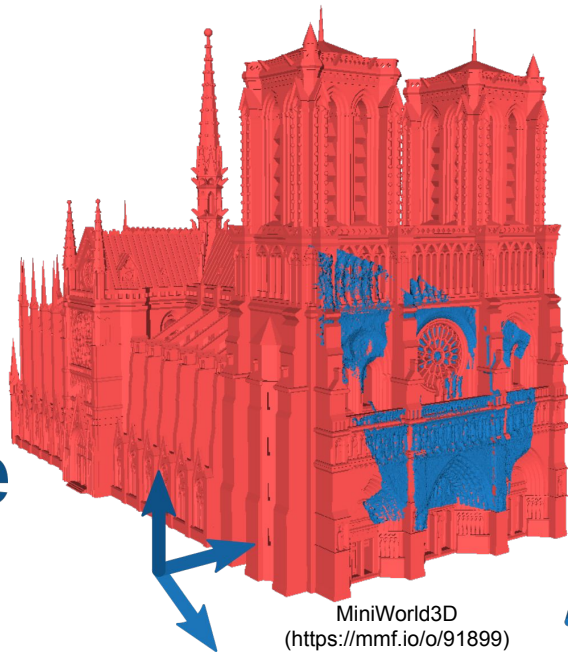
**MVS mesh**

# Evaluation - ground truth pose definition



MiniWorld3D
(https://mmf.io/o/91899)

MiniWorld3D
(https://mmf.io/o/91899)

CZECH INSTITUTE
OF INFORMATICS
ROBOTICS AND
CYBERNETICS
CTU IN PRAGUE

FACULTY
OF ELECTRICAL
ENGINEERING
CTU IN PRAGUE

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA
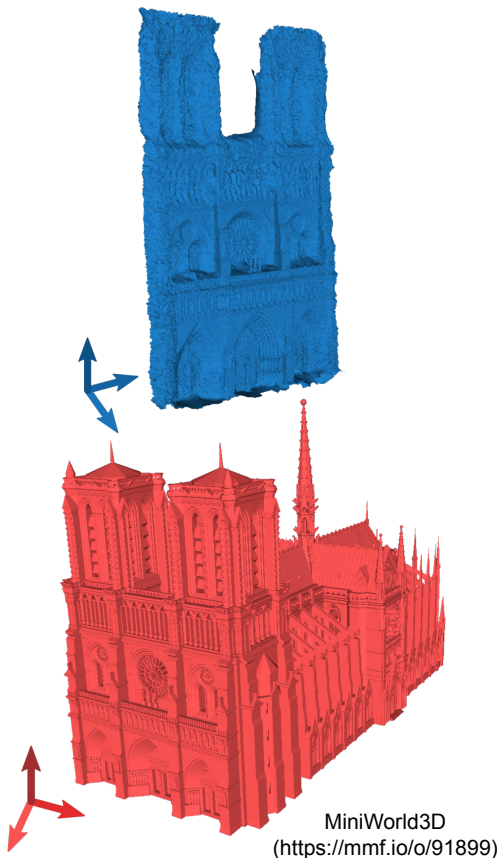
# Evaluation - ground truth pose definition



GT pose

MiniWorld3D
(https://mmf.io/o/91899)
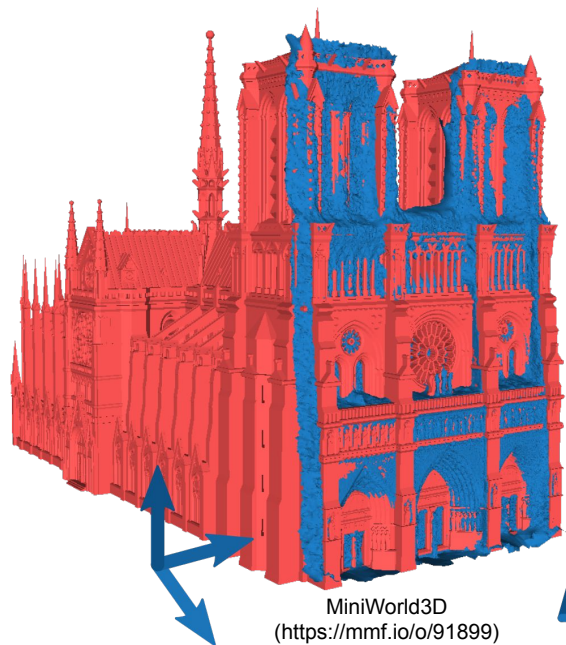
**Global Alignment (GA)**

GT pose

MiniWorld3D
(https://mmf.io/o/91899)

**Local Refinement (LR)**

# Evaluation - metrics



MiniWorld3D
(https://mmf.io/o/91899)

x

y

**estimated pose**   **GT pose**

$RE_{x,y}$

d
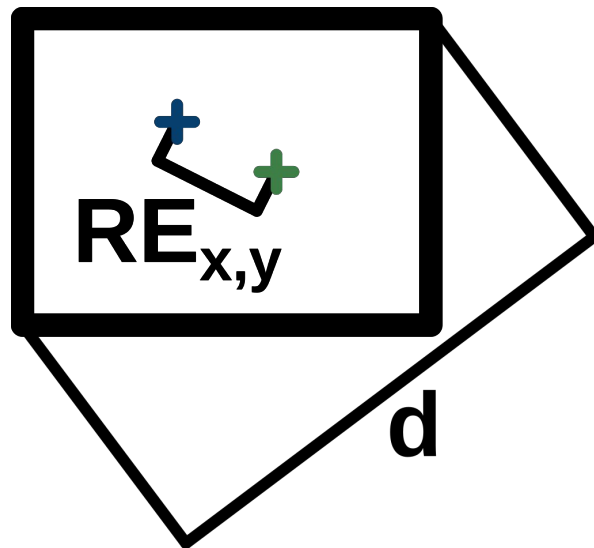
$$DCRE = \{ RE_{x,y} \}_{[8]}$$

$$mean\ DCRE = mean\ (\ RE_{x,y}\ )\ /\ d$$
$$x,y$$

$$max\ DCRE = max\ (\ RE_{x,y}\ )\ /\ d$$
$$x,y$$

[8] Wald, J., Sattler, T., Golodetz, S., Cavallari, T., Tombari, F. "Beyond Controlled Environments: 3D Camera Re-Localization in Changing Indoor Scenes." ECCV, 2020.

# Experiments – localization with rendered images and depth maps
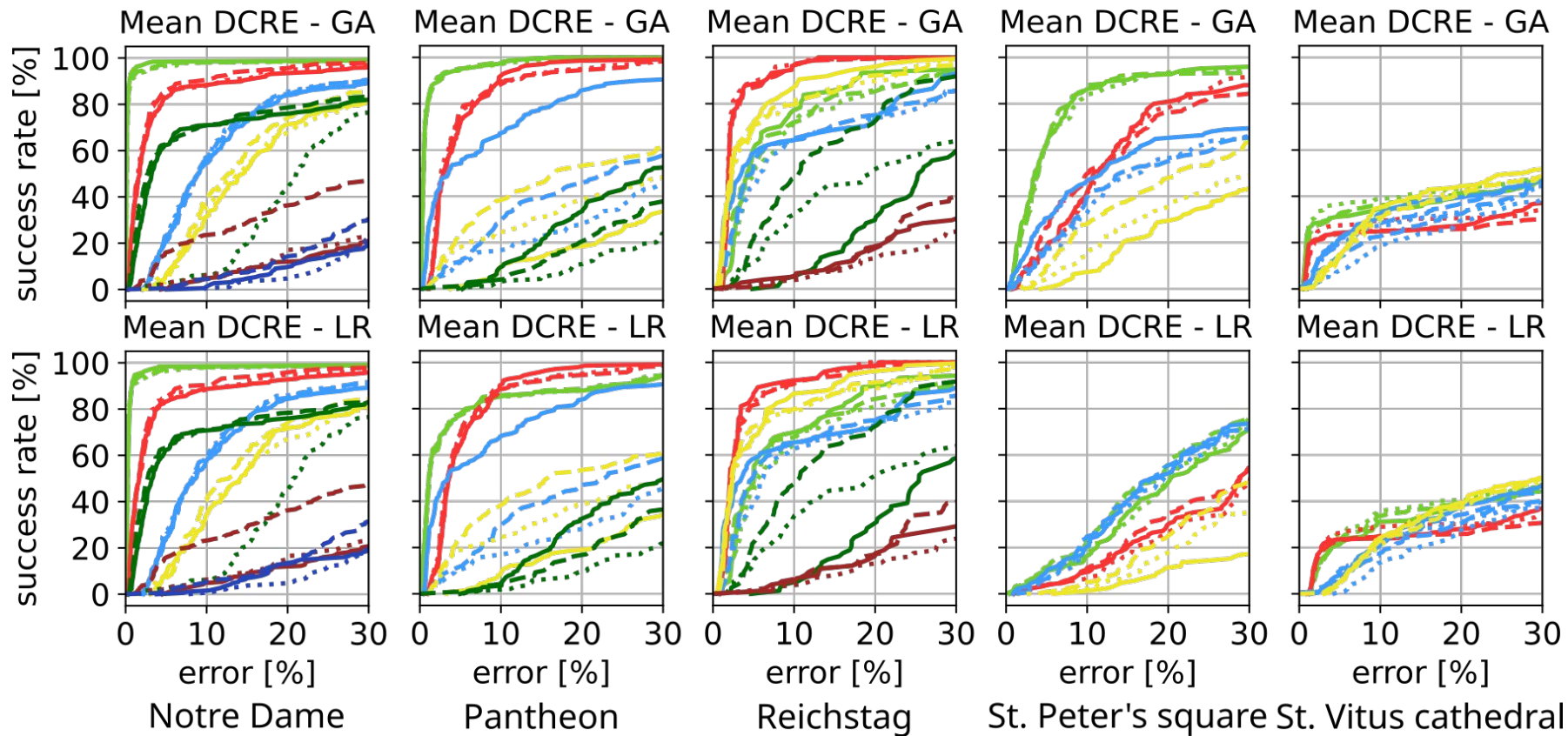
reference
MVS mesh

Fovea
https://skfb.ly/6RZHt

brnipon
https://bit.ly/3CCwTwP

Ultima Ratio
https://bit.ly/3AuN2BK

Adsman007
https://bit.ly/3ASvl0b

Emanuele Viani
https://skfb.ly/EAKB

**rendered images and depth maps**

success rate [%]

mean DCRE / image diagonal [%]

reference
MVS mesh

Fovea
https://skfb.ly/6RZHt

brnipon
https://bit.ly/3CCwTwP

Ultima Ratio
https://bit.ly/3AuN2BK
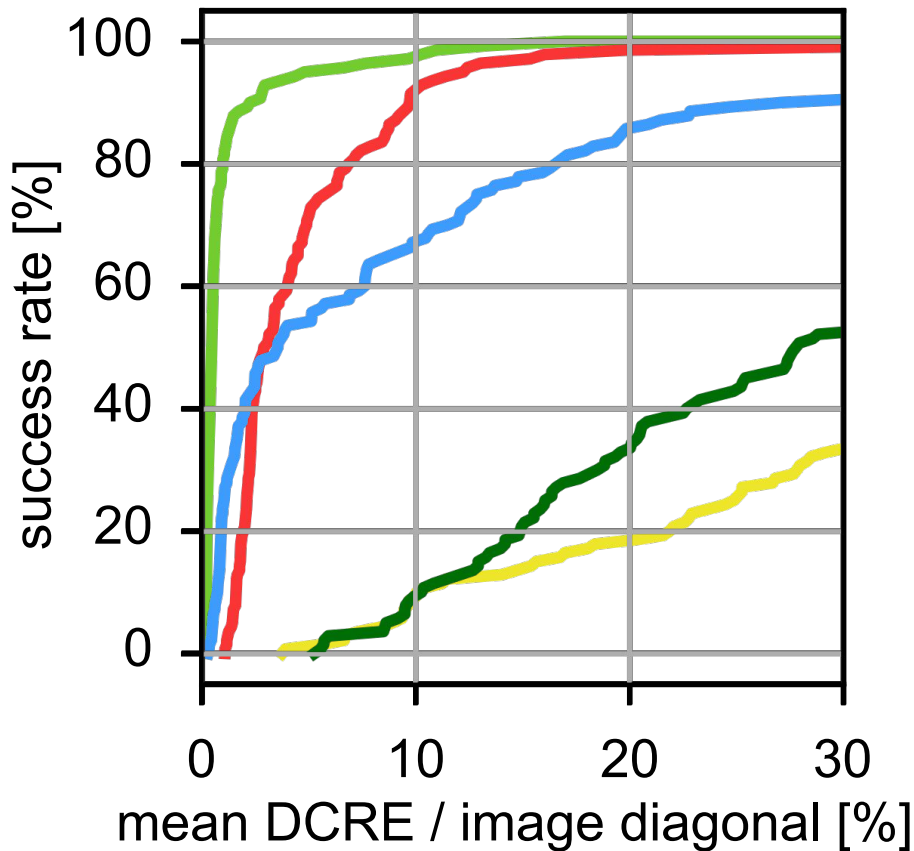
Adsman007
https://bit.ly/3ASvl0b
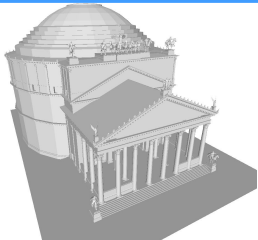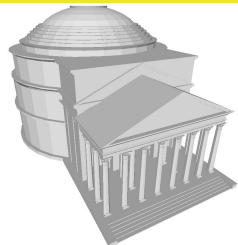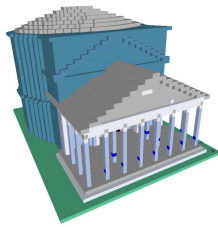
Emanuele Viani
https://skfb.ly/EAKB

**real images + rendered depth maps**

success rate [%]

mean DCRE / image diagonal [%]
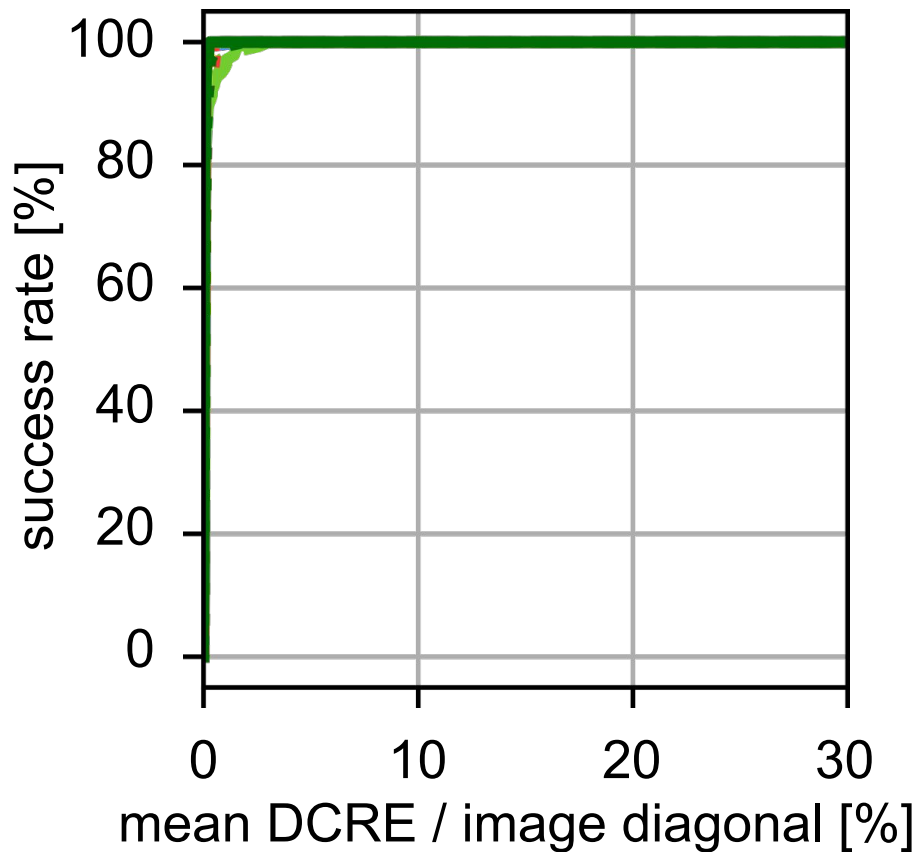
# Experiments – non-uniformly scaled meshes



Miguel Bandera (https://skfb.ly/6QWu7)

| | |
|---|---|
| 80 % | |
| 90 % | |
| 95 % | |
| 96 % | |
| 97 % | |
| 98 % | |
| 99 % | |
| 100 % | |
| 101 % | |
| 102 % | |
| 103 % | |
| 104 % | |
| 105 % | |
| 110 % | |
| 120 % | |
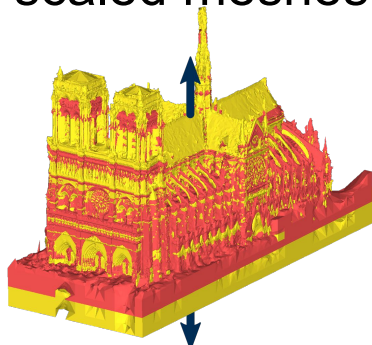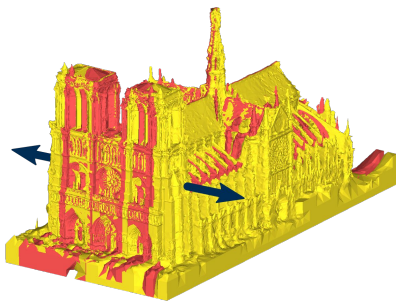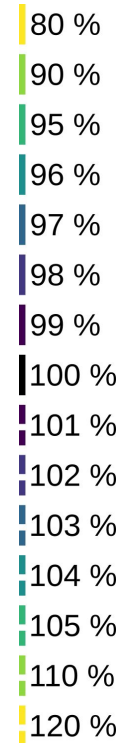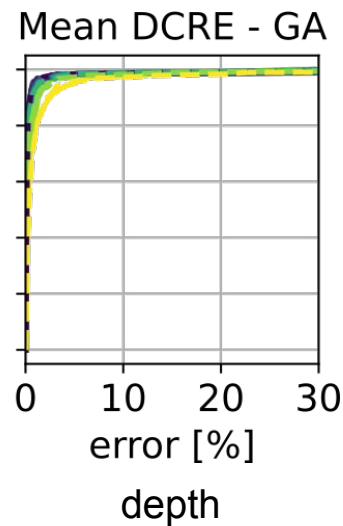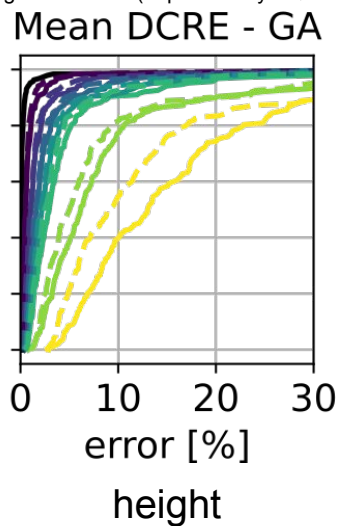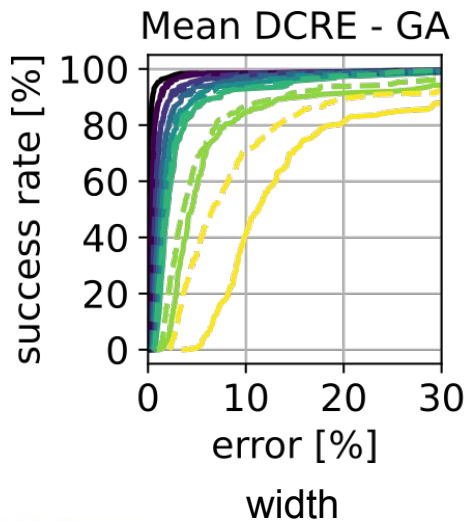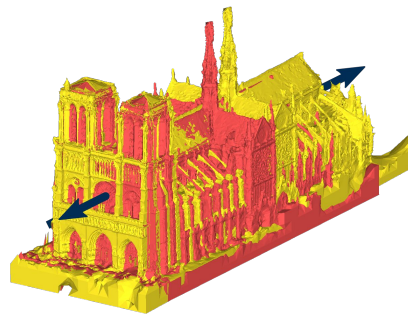
# Experiments – downsampling the geometry and texture



original
Miguel Bandera (https://skfb.ly/6QWu7)
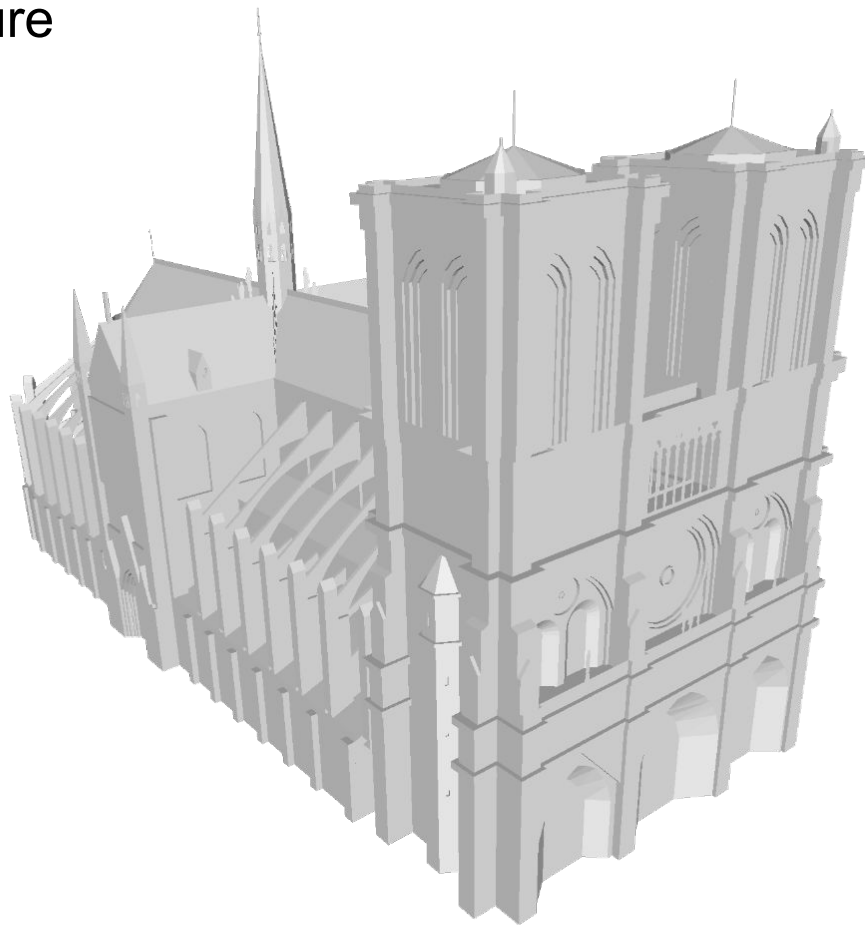
geometry
Mean DCRE - GA
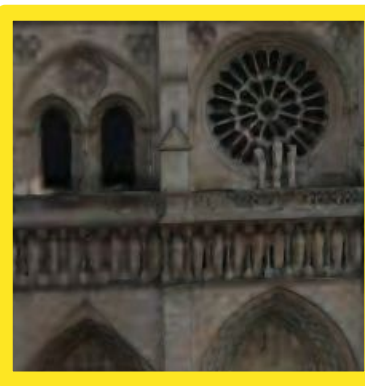
texture
Mean DCRE - GA

both
Mean DCRE - GA

full    1/4    1/16
1/2    1/8
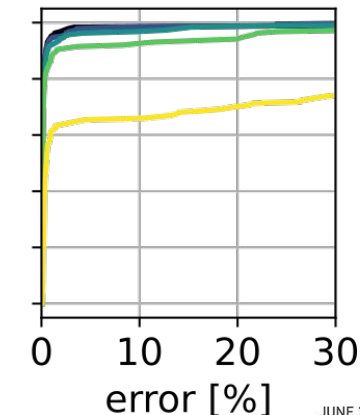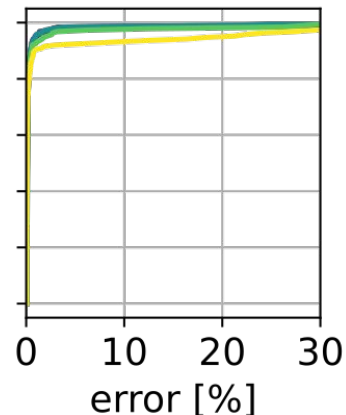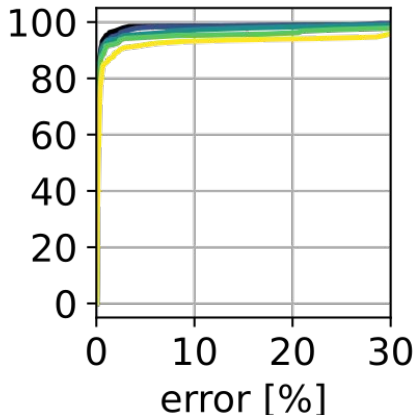
# Data issues - no texture



giotis (https://bit.ly/3QOTQ41)

# Data issues - no texture



Adsman007 (https://bit.ly/3ASvl0b)
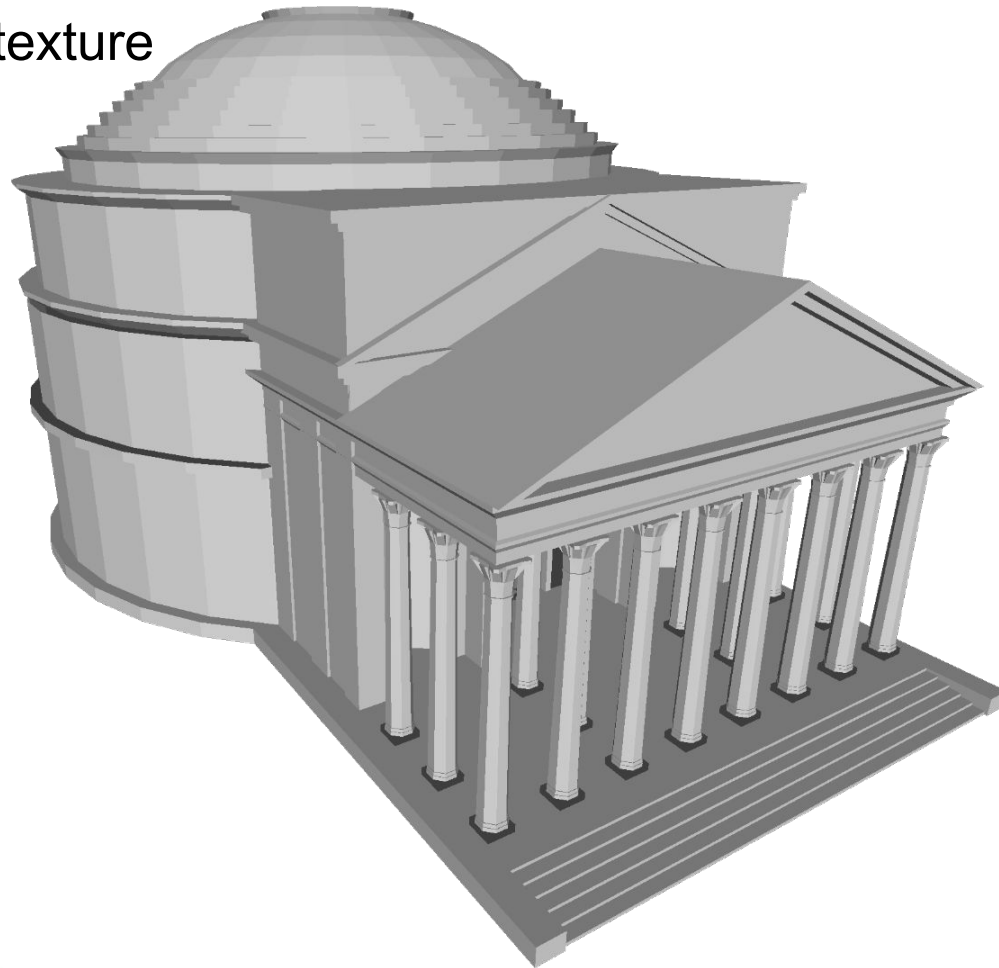
# Data issues - no texture



Klaus T. (https://bit.ly/3cme5qV)
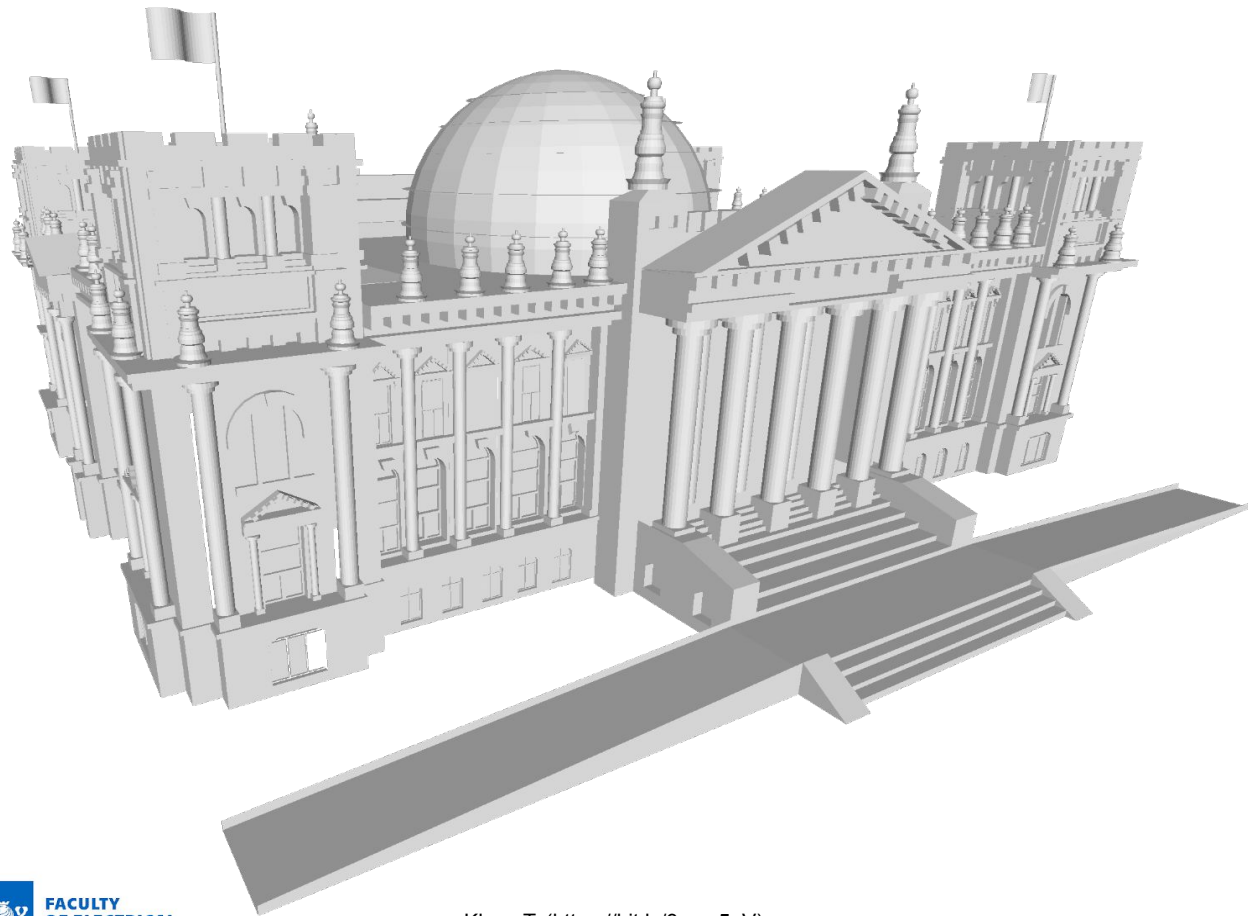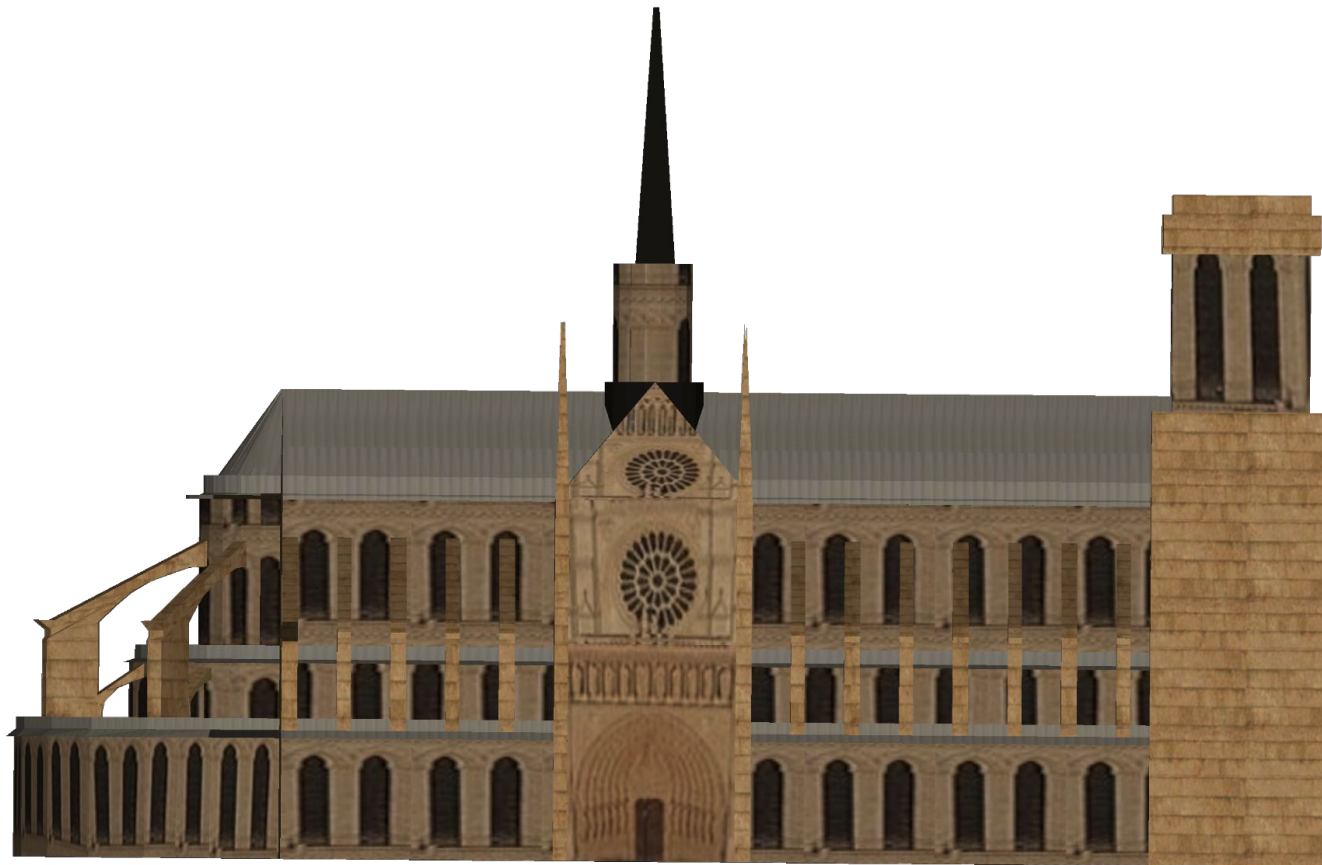
# Data issues - generic / repetitive textures



Little-Goomba (https://bit.ly/3QWeOxY)

# Data issues - generic / repetitive textures



mstochl (https://bit.ly/3RhqEmc)

# Data issues - generic / repetitive textures



Pera (https://bit.ly/3Tf7Bum)

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA
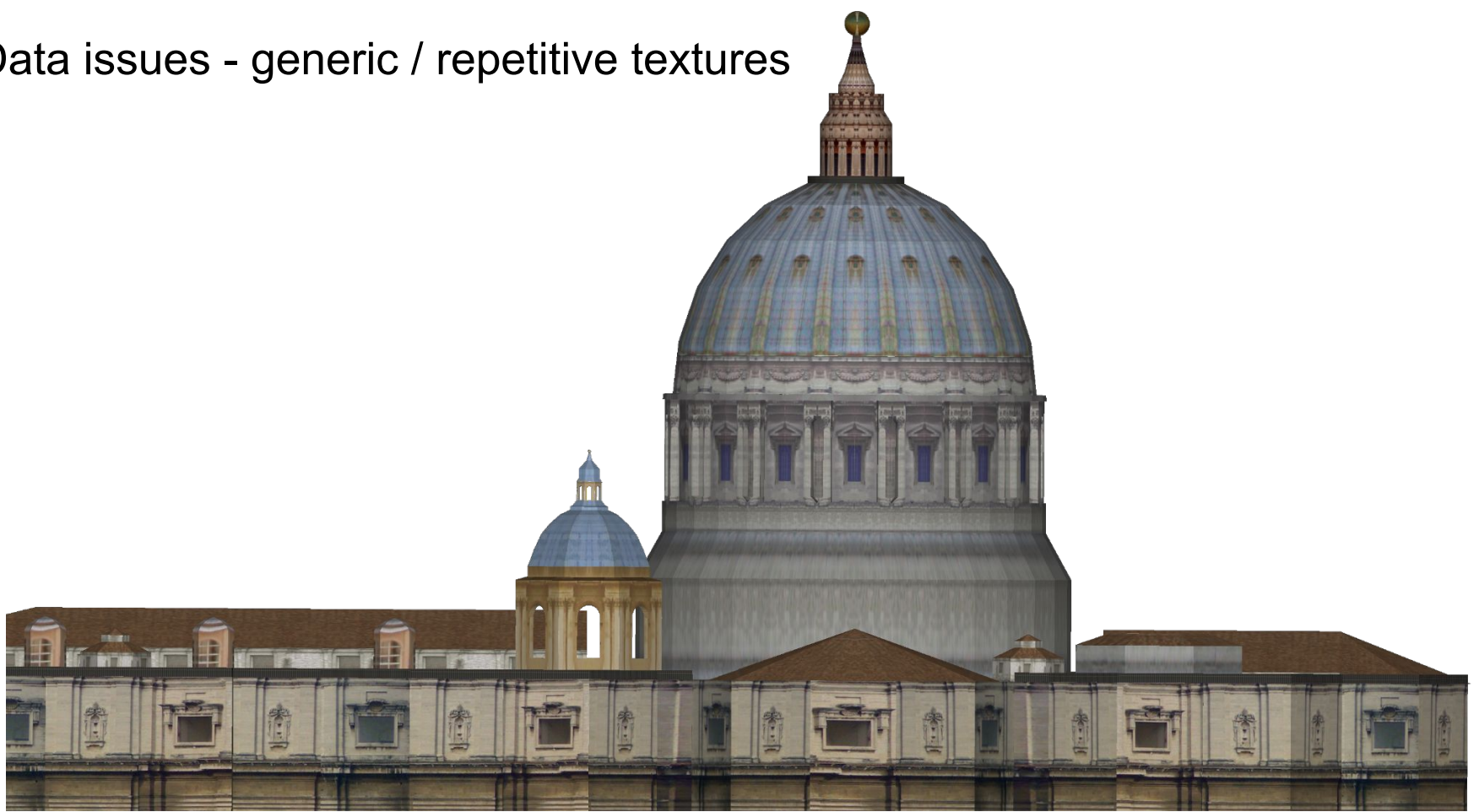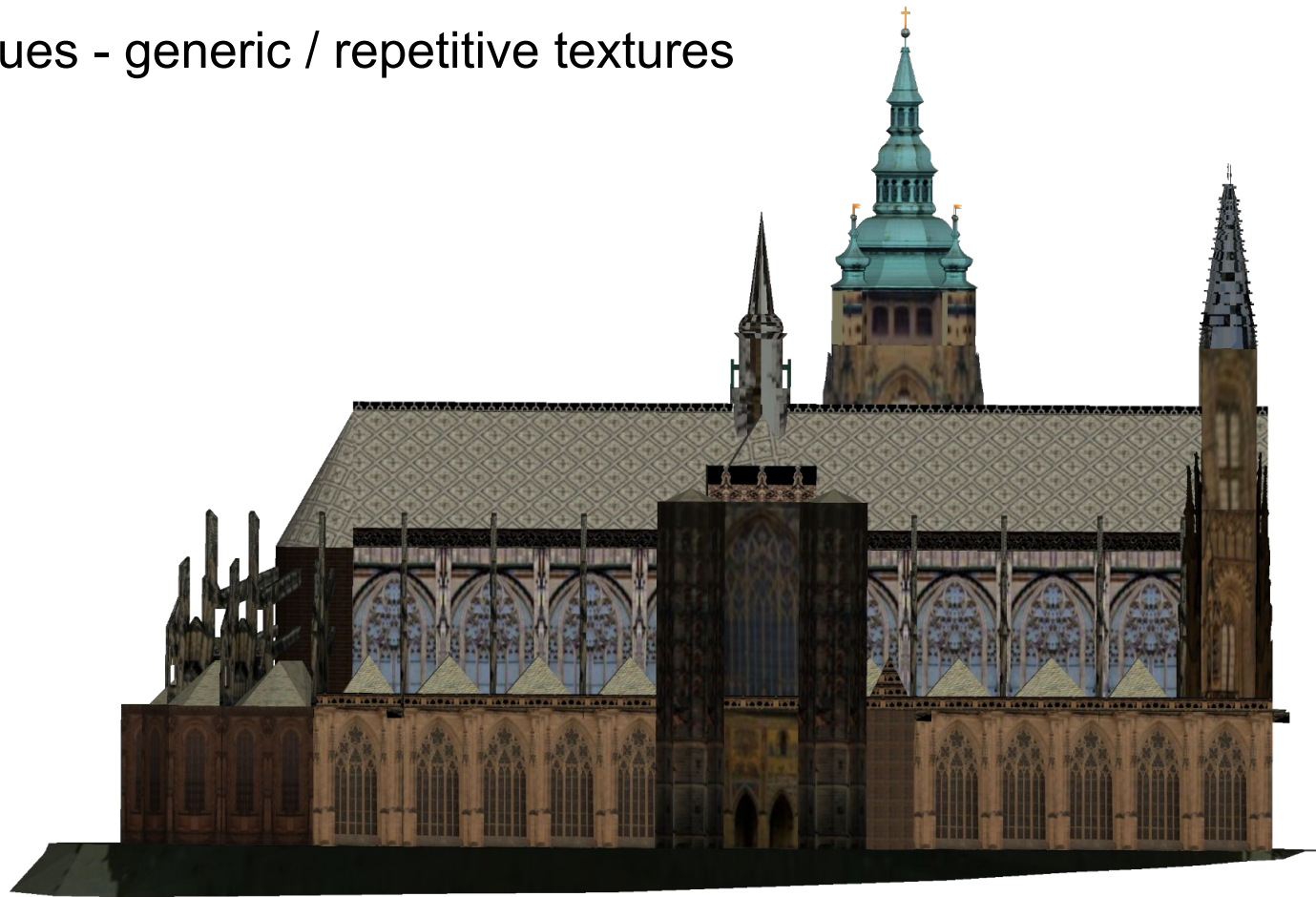
# Data issues - coarse geometry



Jul (https://bit.ly/3ThrhOt)
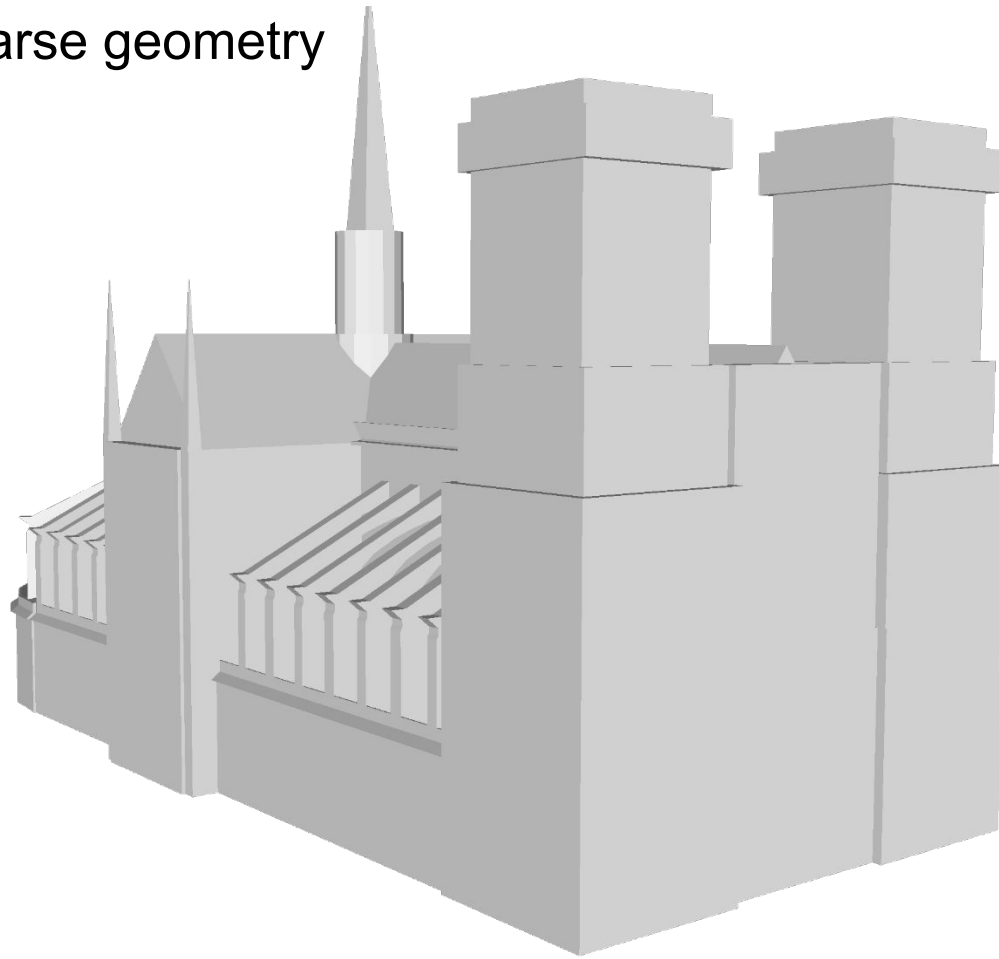
# Data issues - coarse geometry



Emanuele Viani (https://skfb.ly/EAKB)
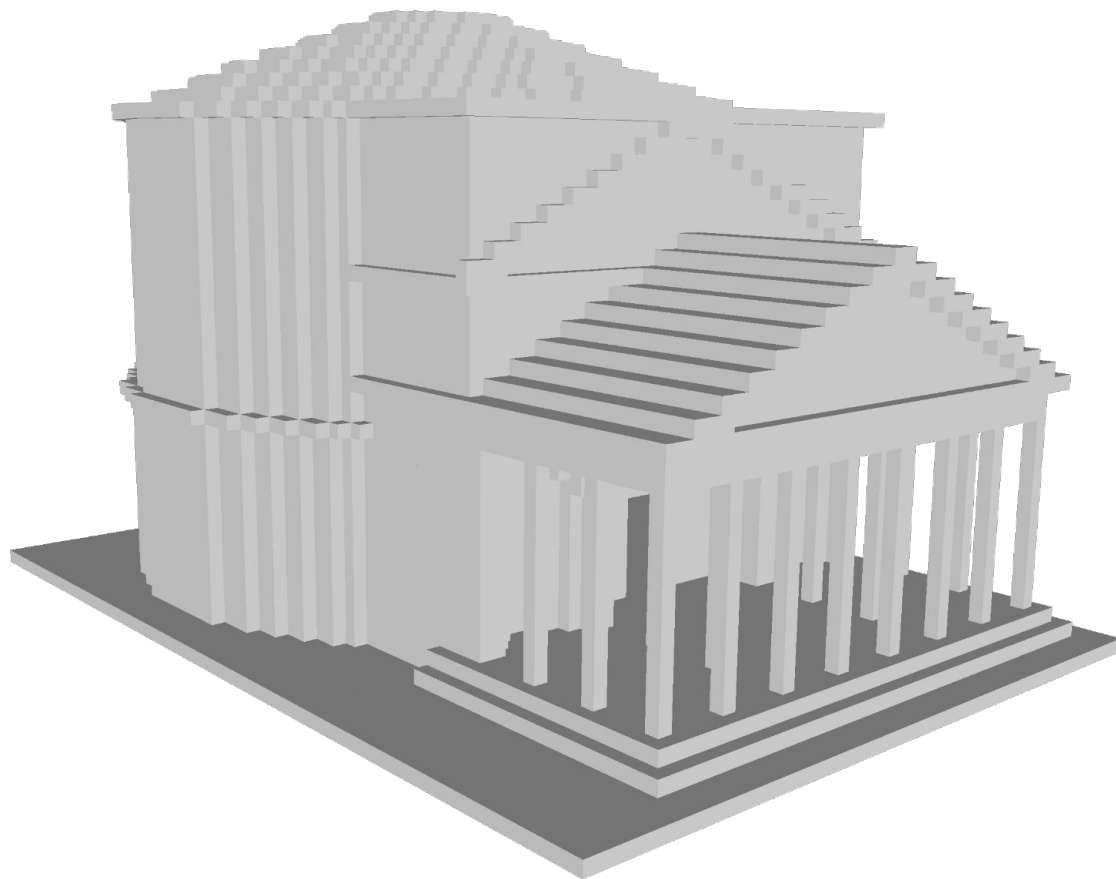
# Data issues - non-existent geometry



Ultima Ratio (https://bit.ly/3AuN2BK)
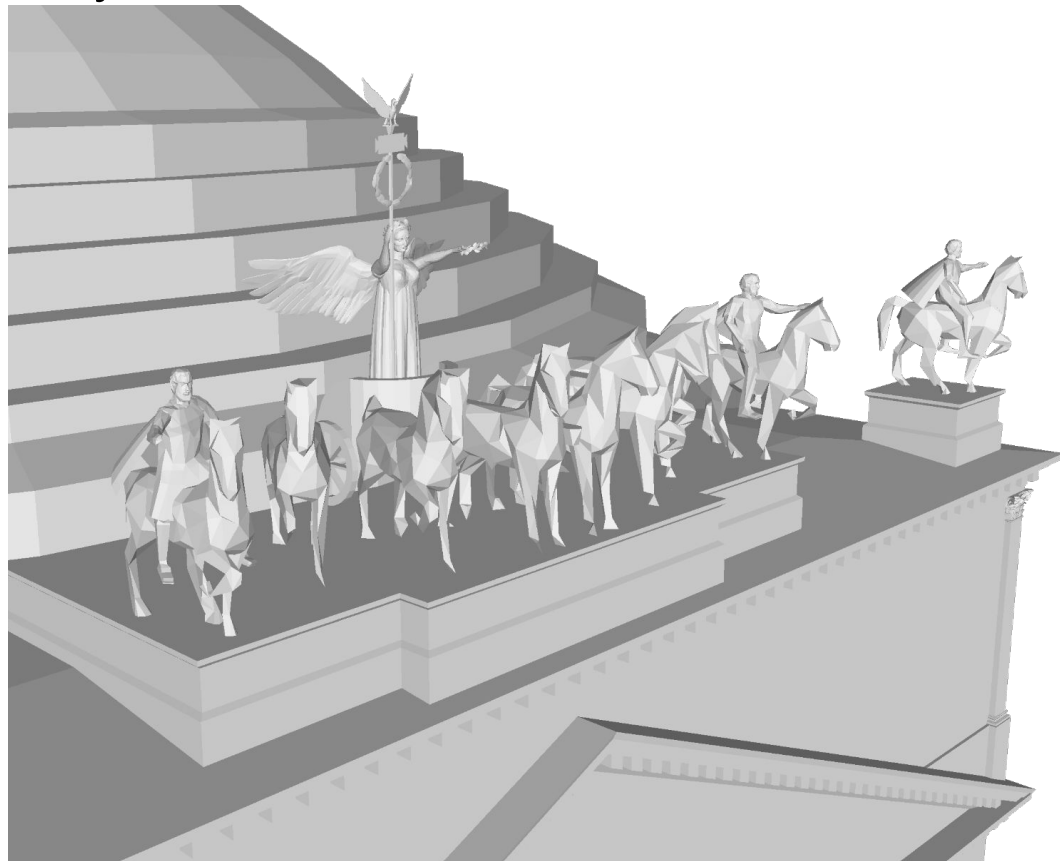
# Data issues - non-existent geometry



Emperor Heer 99 (https://bit.ly/3ctJvvp)
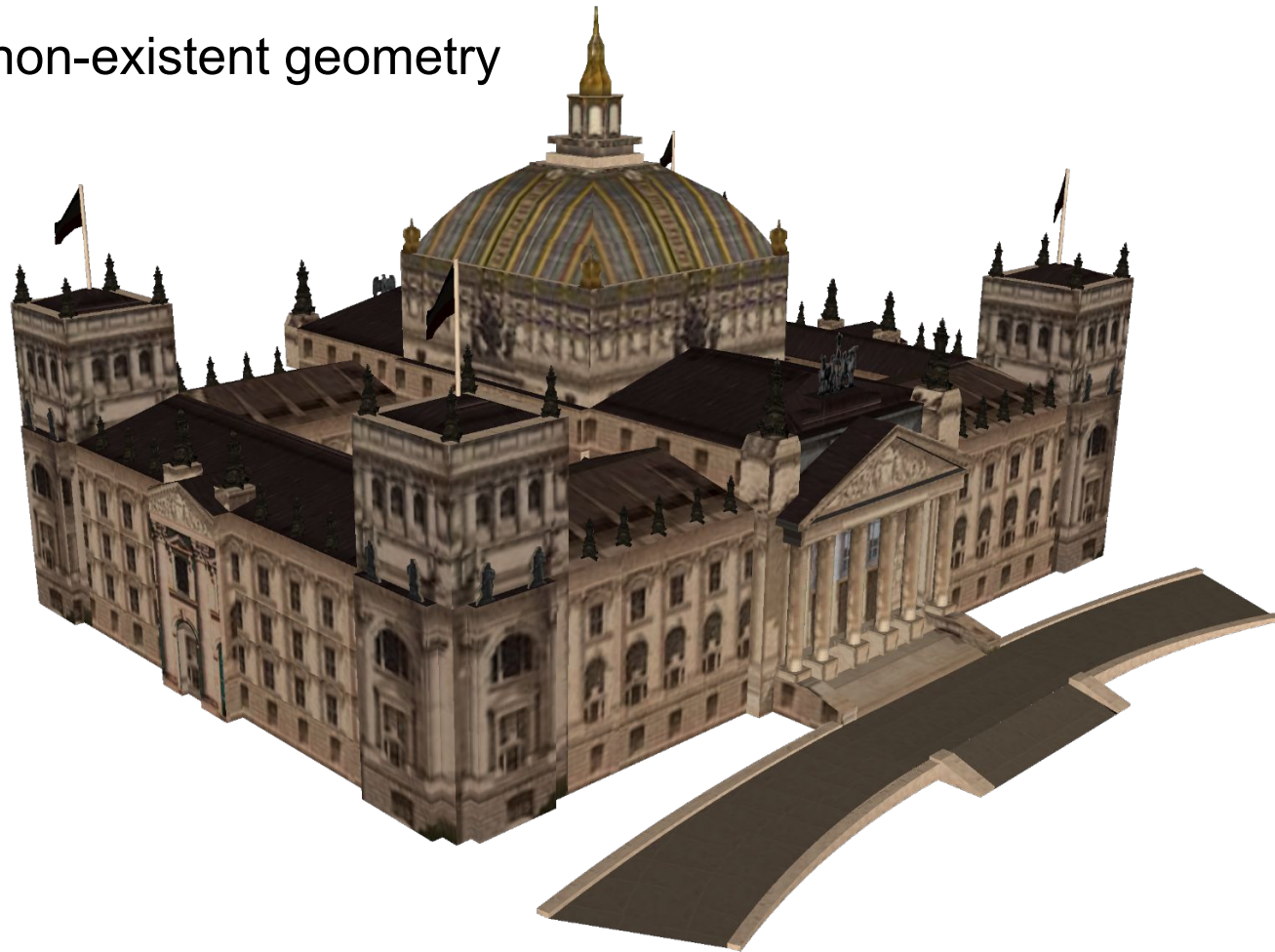
v-pnk.github.io/cadloc

Benchmark for visual localization on 3D models from the Internet

- dataset with gathered 3D models
- multiple scenes
- multiple model quality and detail levels
- color / raw geometry
- rendering and evaluation scripts

**WED-PM-078**