



# XR・メタバース最前線 — AI・3D再構成・デジタルツインが 切り拓く新たな地平

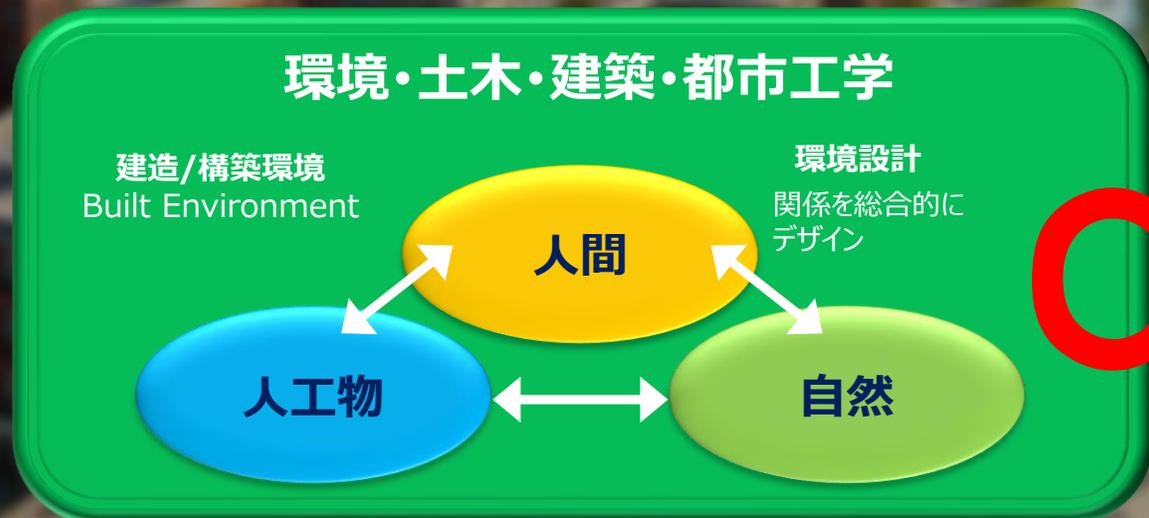
大阪大学 大学院工学研究科

教授 福田知弘



# 環境設計情報学領域: 研究テーマ

- Society 5.0すなわち「超スマート社会」の構築を目指して、人間・人工物・自然などの関係を総合的に設計する環境デザインの方法論を構築すると共に、情報通信技術（ICT）の高度利用による環境デザインシステムの開発および総合工学的な環境情報学の研究・教育を行っています。
- 環境・土木・建築・都市工学におけるICTに関する研究、特に、AI、デジタルツイン、拡張・複合現実とバーチャルリアリティ、3次元計測とプロダクトモデルなどの技術に関する研究、ならびに、それらを共生環境デザインに応用する研究を行います。



# 建造環境の変化

大中遺跡（兵庫県 播磨町）

弥生時代終末（約1900年前）から古墳時代初頭



第一土器群下住居（弥生時代後期）  
（この復元家は遺構の状況から推定したもので元の位置を移動させている。  
この建物は 第一土器群の時代（弥生時代）の推定された下から発見された  
住居跡を推定復元したものである。  
やや長方形の平面をもち、短辺に中央より高い土床がつけられている。  
短辺2本で 直線線が交差しているが、このように簡単な構造の建  
物は他にほとんど見つかっていない。また、このように見られるものは、  
なお入口は西側、短辺に入っているから、正しようは平面的のもの  
で裏へ側へ入る ものである。この建物は入口の位置  
が、さきついでにわかったことわかる。



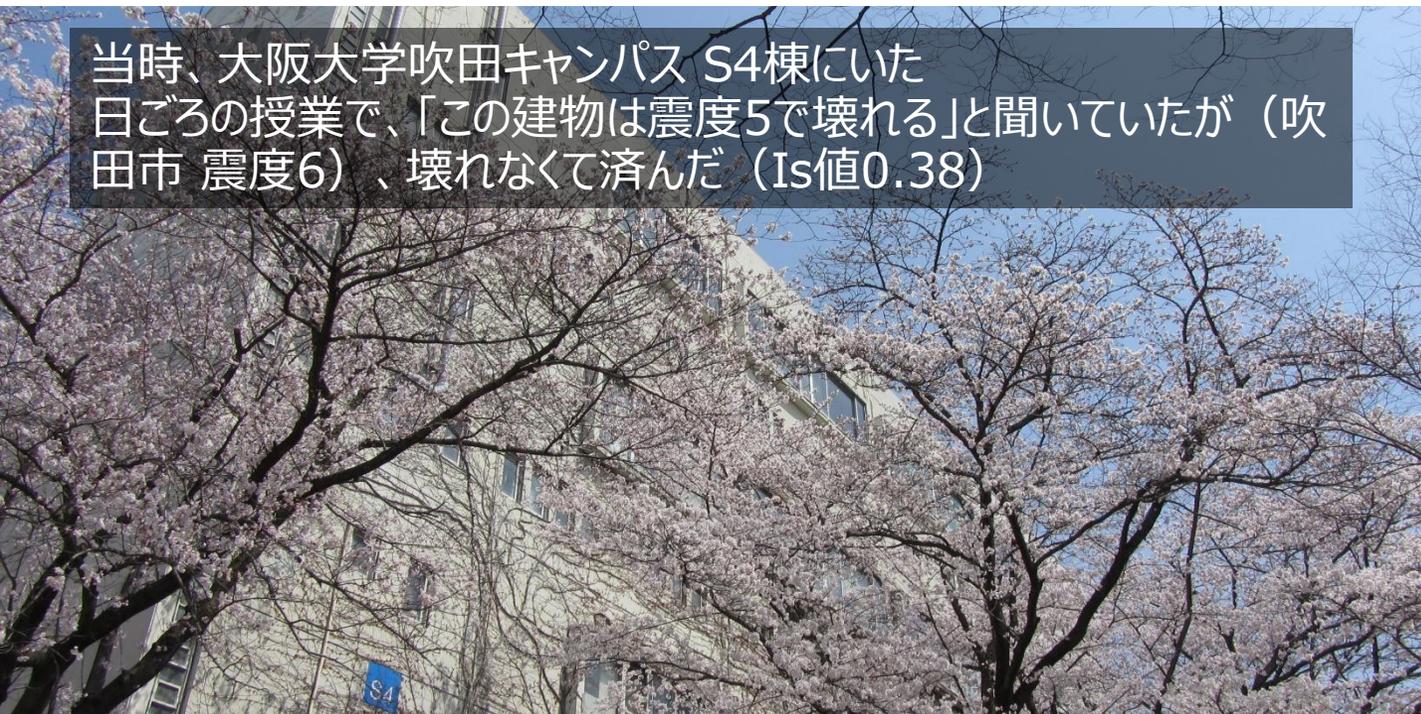
Photo: 神戸市

# 阪神淡路大震災 1995年1月17日5時46分

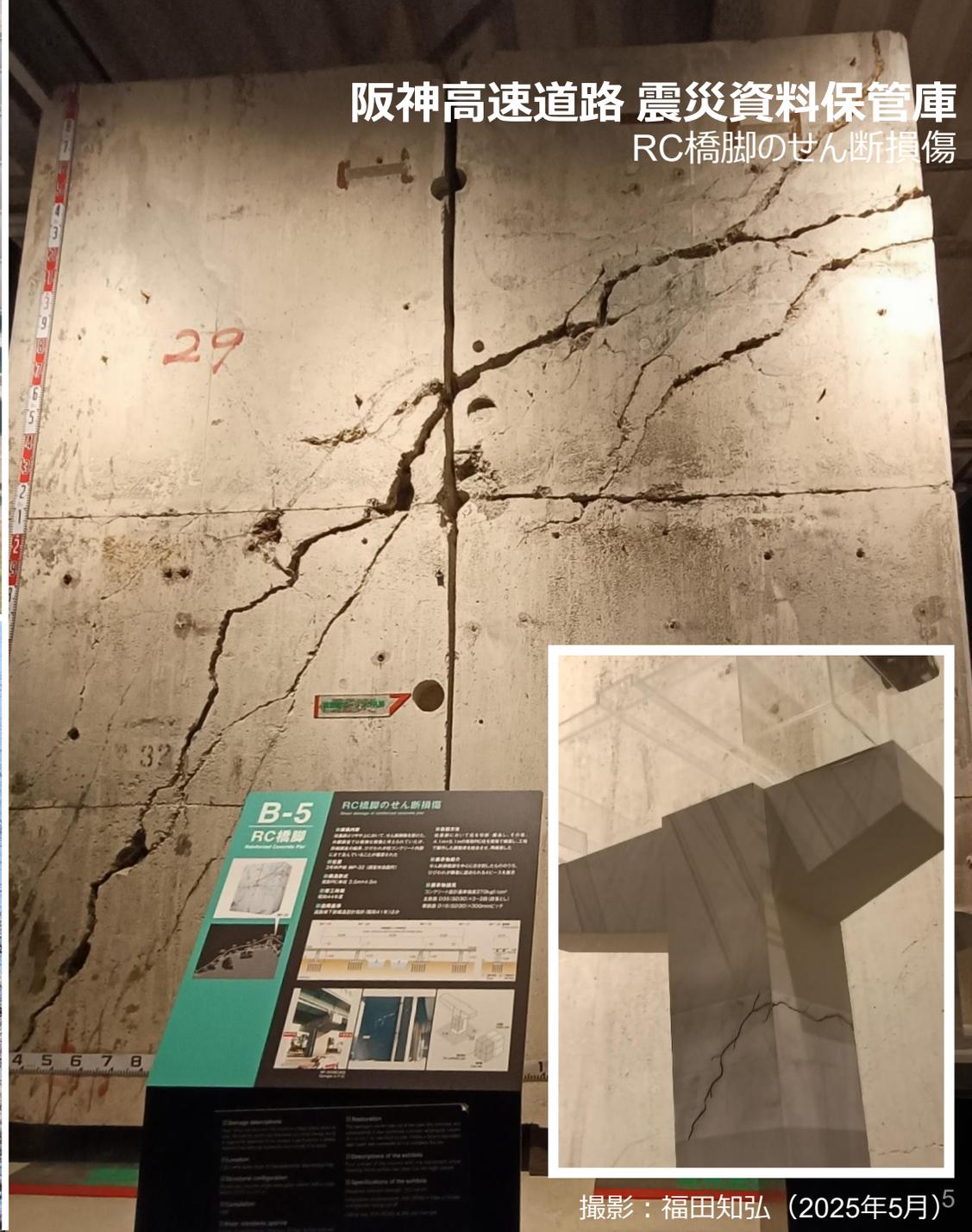


写真提供：神戸市

当時、大阪大学吹田キャンパス S4棟にいた  
日ごろの授業で、「この建物は震度5で壊れる」と聞いていたが（吹  
田市 震度6）、壊れなくて済んだ（Is値0.38）



## 阪神高速道路 震災資料保管庫 RC橋脚のせん断損傷



撮影：福田知弘（2025年5月）<sup>5</sup>

# 建造環境の ライフサイクルと 利害関係者



リノベーション/  
解体撤去



運営管理

施工

施主

国・地方自治体

民間デベロッパー

投資家

近隣住民

**BIM**  
(プロダクトモデル)

非専門家

設計者  
エンジニア

専門家

利用者  
一般市民  
ジャーナリスト

ソフトウェアベンダー

施工者  
管理者

利害関係者

計画

設計

ステイクホルダーエンゲージメント  
参加型まちづくり／スマートシティ



デジタルサイネージ  
メディアファサード

# XR (VR・AR・MR・DR), メタバース, デジタルツイン



バーチャルモデル



センシング ↑

デジタルツイン (Digital twin)

シミュレーション

↑

現実世界・  
架空世界

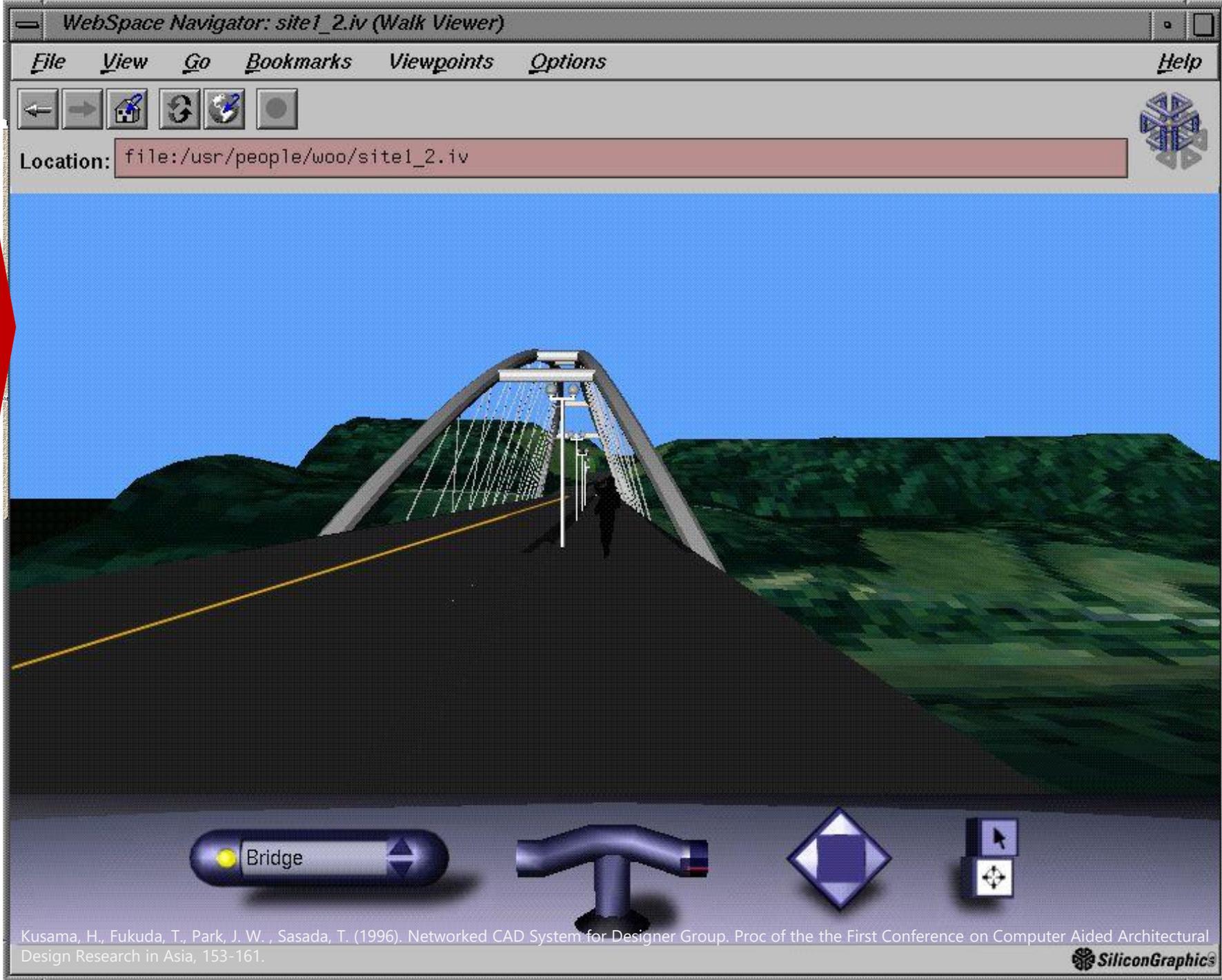
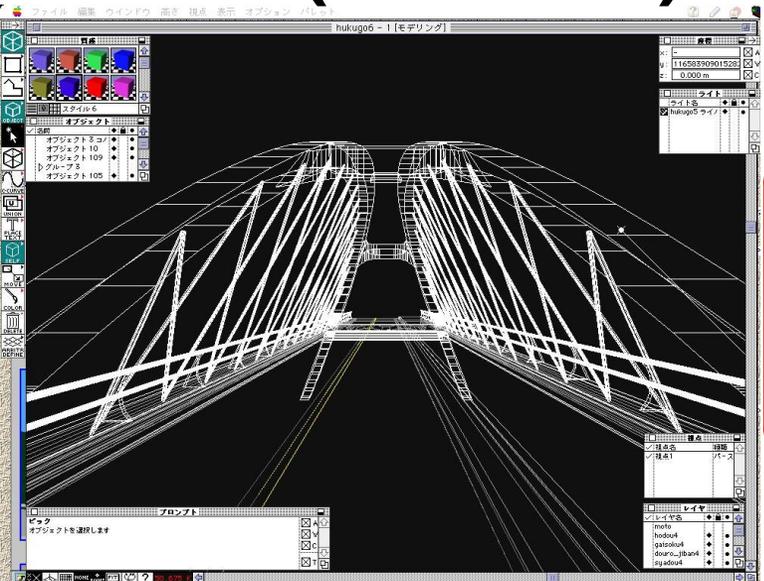
MRメタバース

メタバース (Metaverse)

# VR



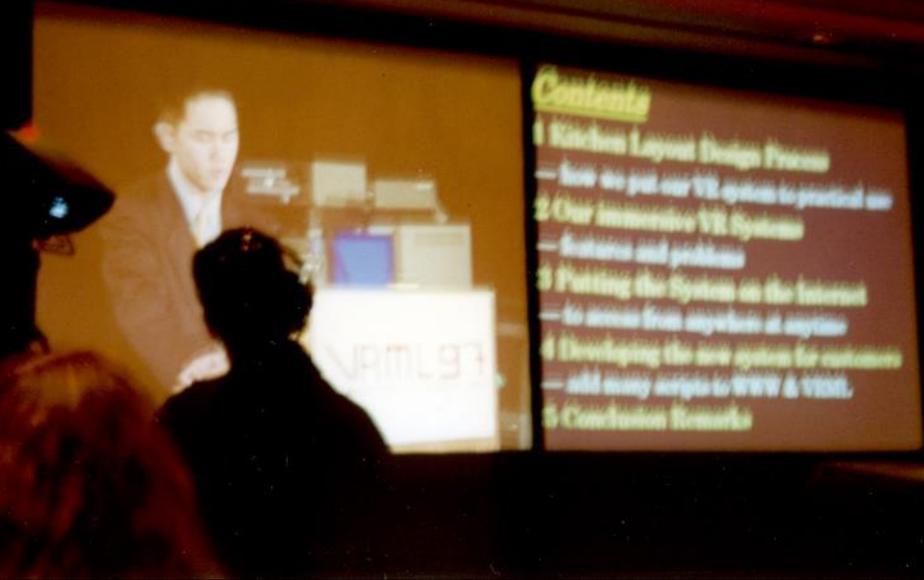
# VR(1995)



Kusama, H., Fukuda, T., Park, J. W., Sasada, T. (1996). Networked CAD System for Designer Group. Proc of the the First Conference on Computer Aided Architectural Design Research in Asia, 153-161.

# VRML97 国際会議 (Monterey, CA: 1997)

## 2<sup>nd</sup> Symposium on the Virtual Reality Modeling Language



**CPU: Pentium-Pro 200 MHz**  
**OS: WinNT 3.5.1**  
**RAM: 64 MB**  
**HDD: 2 GB**  
**GPU: 登場前 (GeForce 256: 1999.10)**



menu Reset

自由飛行

自由歩行

編集

現状

計画1 昼夜

計画2 昼夜

検討項目

樹木

- ケヤキ1 モクレン1
- ケヤキ2 モクレン2
- ケヤキ3 モクレン3

ベンチ

- ベンチ1
- ベンチ2

交番

街灯

電柱

標識

ゴミ箱

電話ボックス

自転車

分電盤

人

高松4町パティオ (2005-06)  
CAADRIA 2007, Chiang Mai



VR安土城 近江八幡市 (2013)  
CAADFutures 2015, Sao Paulo



回廊の家VR 潮来市 (2015)  
eCAADe 2015, Vienna



図書交流広場

境港市民交流センターVR 境港市 (2019)  
eCAADe 2021, Novi Sad (Online)



高松4町パティオ (2007)



VR安土城 (2013)



回廊の家VR 潮来市 (2015)



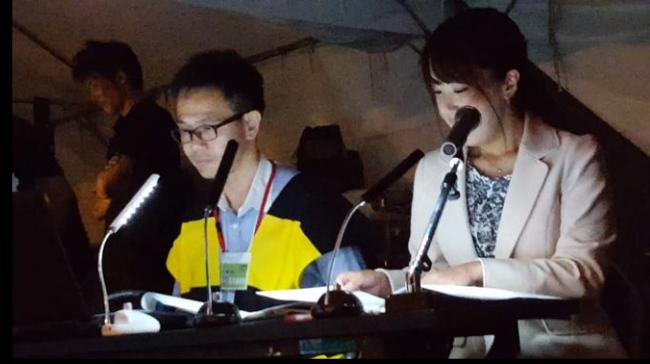
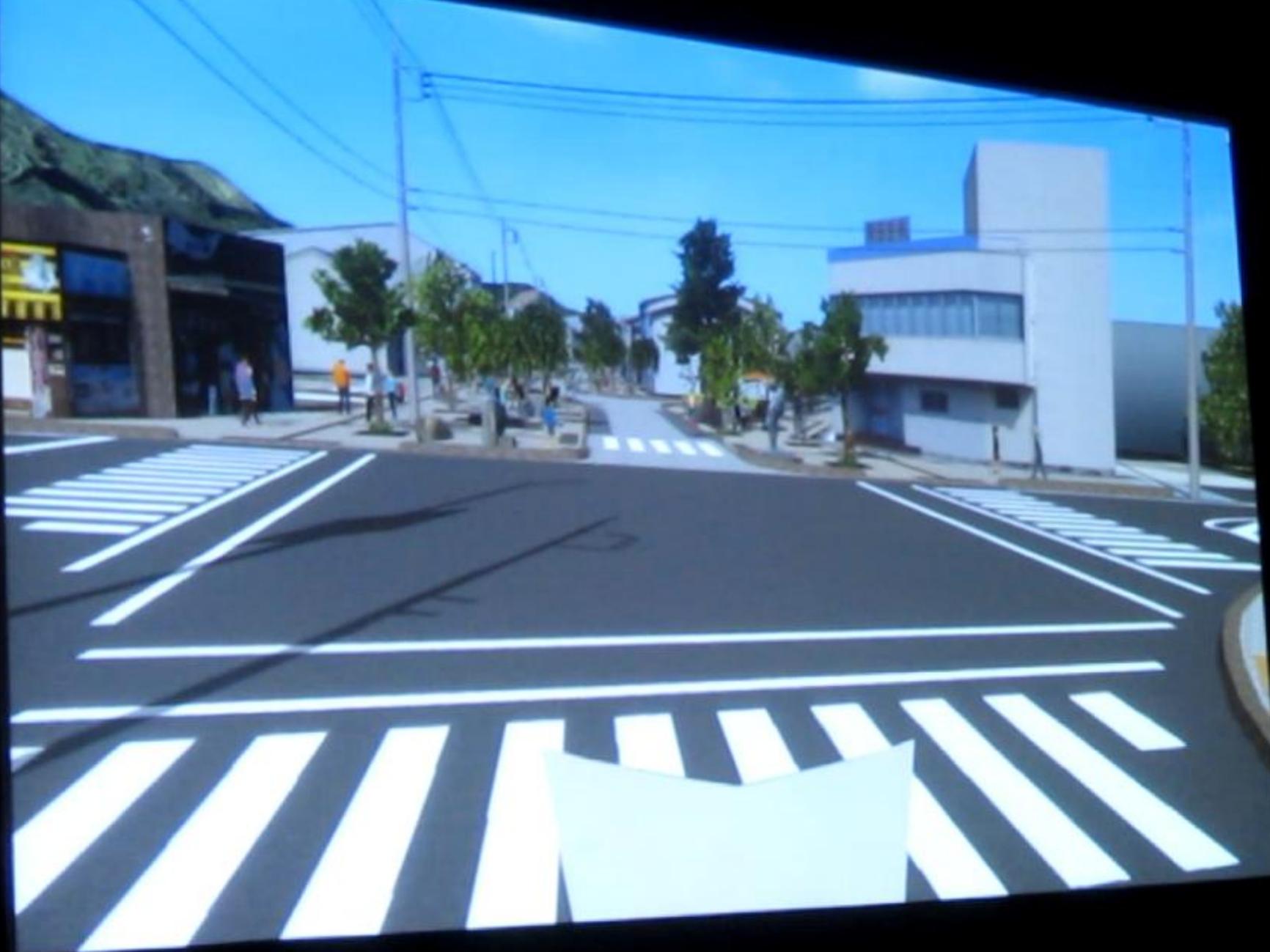
設計 & 写真: atelier DoN

境港市民交流センターVR (2019)



# 水木しげるロードリニュアルVR 境港市 (2016)

CAADFutures 2017, Istanbul



# MAR



# MRシステム構築 (2005)



# 高松4町パティオ (2006)

eCAADe 2006, Volos



# 回廊の家AR (2015)

eCAADe 2015, Vienna



# 丹後国分寺五重塔AR (2022)



# 鳥取城 三階櫓復元 AR (2023)



# 温熱環境検討のための BIM・CFD・MRの統合 (2017)

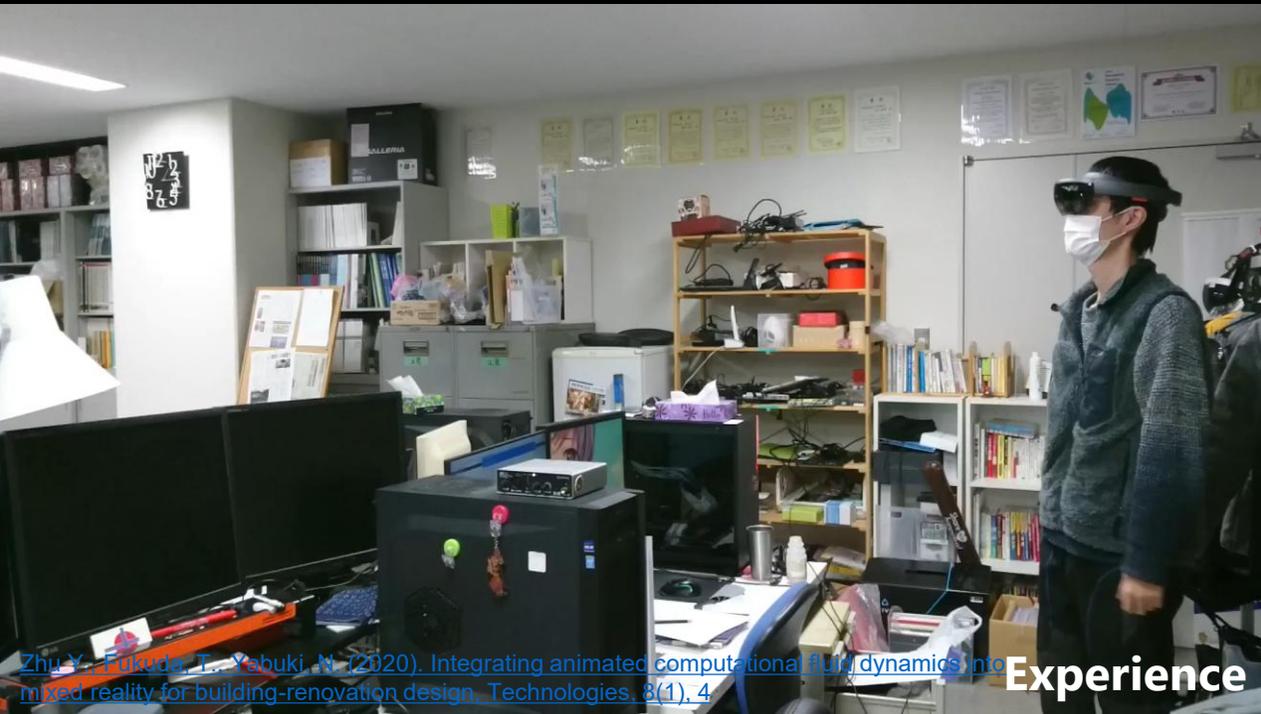


HMD view

# テレプレゼンスに向けた 点群ストリーミングのMR表示 (2017)



HMD view



Experience



Remote site



Experience



# MRRDR

# AI

# AI-powered Mixed Reality (2015—)

- 通常のMR/DR（複合現実/隠消現実）は現実世界と仮想モデルを重ね合わせ、将来景観を表現する [1]。
- しかし、現実世界と仮想モデルの相互作用などは限られ、不正確である。
- AIは、現実世界の何かを検出できる。新しい世界を生成することもできる。
- **MR/DRとAI（機械学習/深層学習）を統合して、MR/DRを高度化できないか？**



# 建造物のMR/DR (複合/隠消現実) 2015-17



ライブ映像  
大阪大学 吹田キャンパス ポプラ通り福利会館



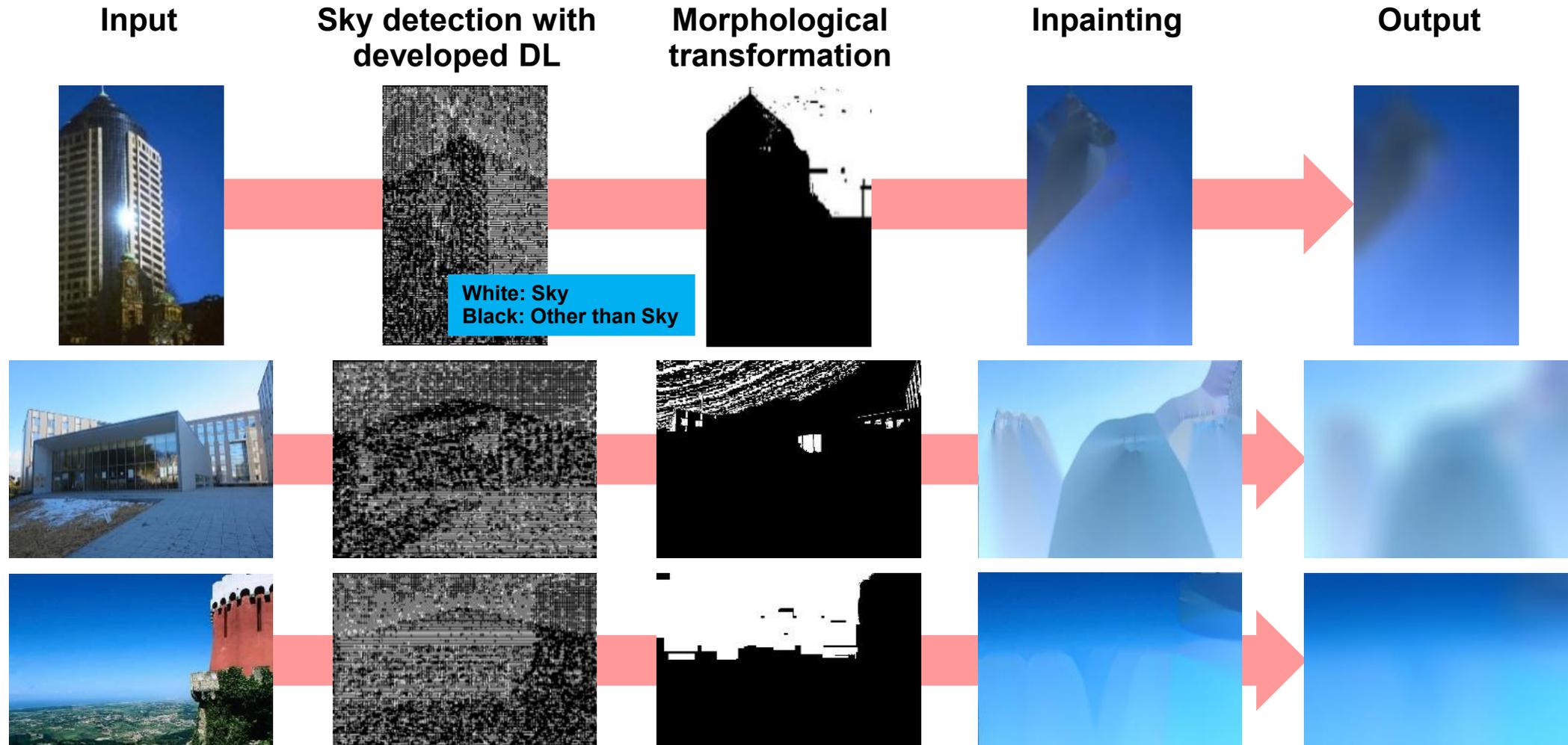
MR景観シミュレーション (不正確)



MR・DR景観シミュレーション (解決策)

# DL（深層学習）による空検出（2016）

- DR（隠消現実）によって生成された背景は、現在の空情報からの手がかりで補完したい[1]。
- SegNetを参考に空検出モデルを開発する[2]。



[1] Fukuda, T., Kuwamuro, Y., Yabuki, N. (2017). Optical Integrity of Diminished Reality Using Deep Learning, Proceedings of eCAADe 2017, 1, 241 – 250

[2] Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12), pp. 2481–2495, 7803544

# ビル取り壊し後の景観を即再現 (2021)

- セマンティックセグメンテーション (建物検出) とGAN (建物除去領域のインペインティング) の統合



Current Scene



Future Scene



Semantic Segmentation



Mask

Generative adversarial network (GAN: 敵対的生成ネットワーク)



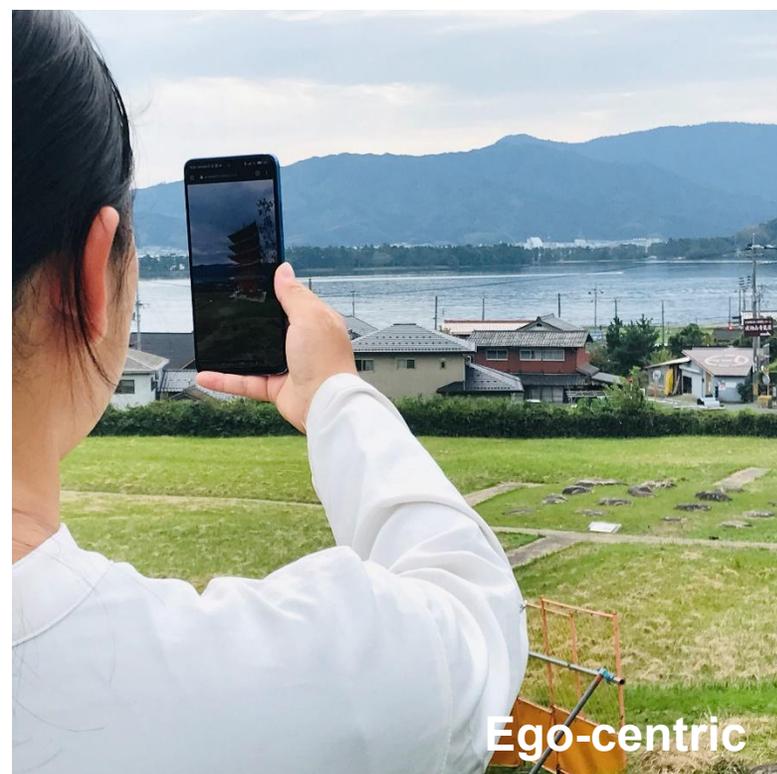
An aerial view of a city with several drones flying in the sky. The image is overlaid with a semi-transparent grid of data visualizations, including bar charts, circular gauges, and text labels like '3D MODELS', 'RESTRICTIVE', and 'RETAVERSE'. The text 'MR AI' is written in a large, bold, white sans-serif font across the top, and 'Drone' is written in a large, white, rounded sans-serif font across the bottom.

# MR AI

# Drone

# Drone-enabled Mixed Reality (2018-)

- MR（複合現実感）は通常、ユーザの視点近傍からの体験に限定される。
- ユーザ近傍から離れた上空など、より自由な視点からMR体験できないか？





# 都市デジタルツインによるMRオクルージョン処理



現状映像



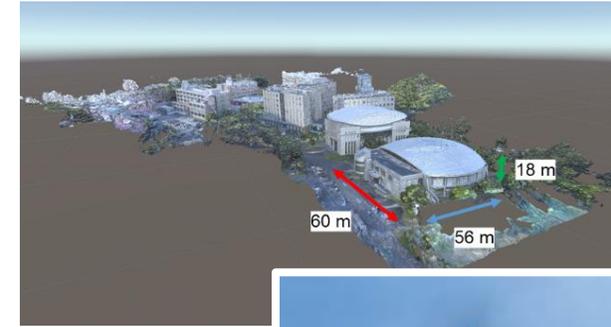
3D 仮想世界（既存都市モデルと設計モデル）



オクルージョン処理なしMR



オクルージョン処理付きMR

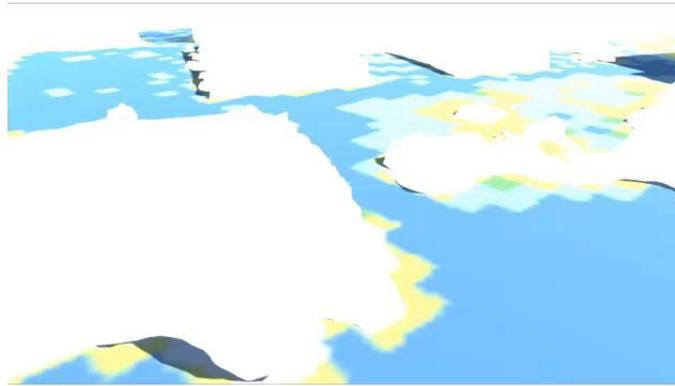


Kikuchi, N., Fukuda, T., Yabuki, N. (2022). Future landscape visualization using a city digital twin: Integration of augmented reality and drones with implementation of 3D model-based occlusion handling. *Journal of Computational Design and Engineering*, 9 (2), 837 – 856

# ドローン視点からのMR: 都市洪水の可視化



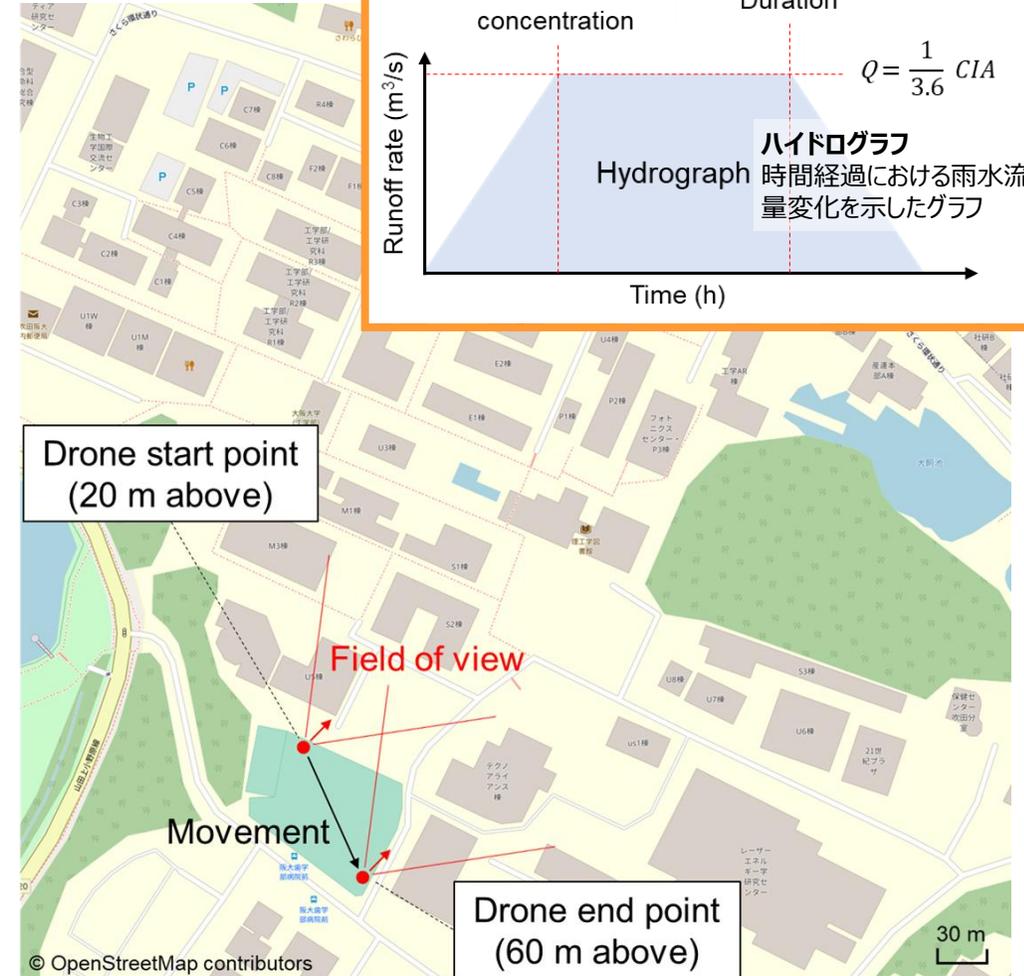
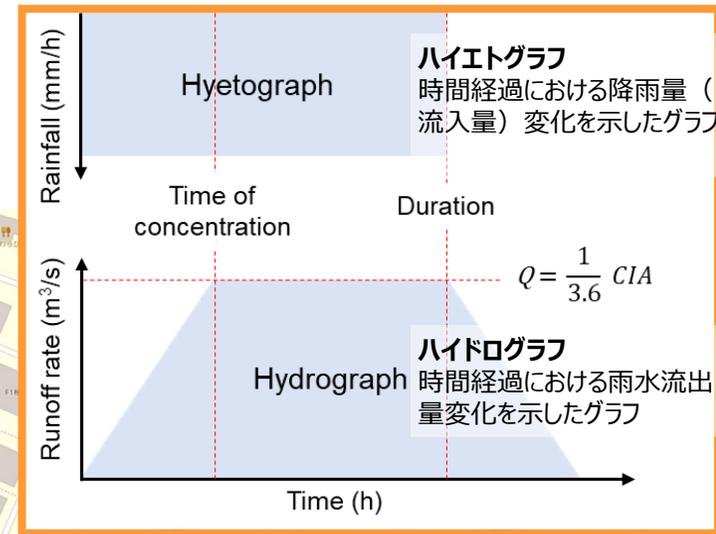
現状映像



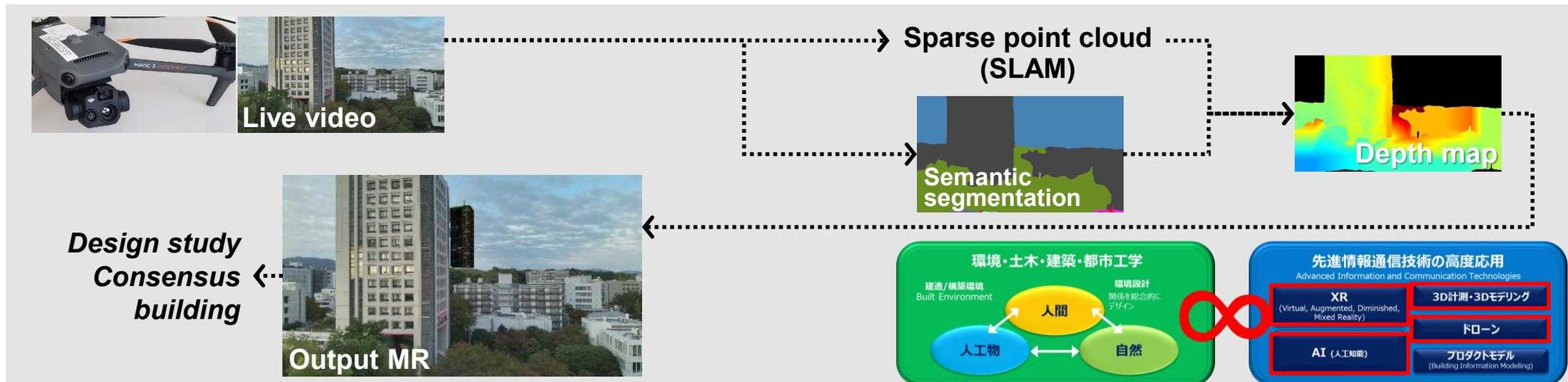
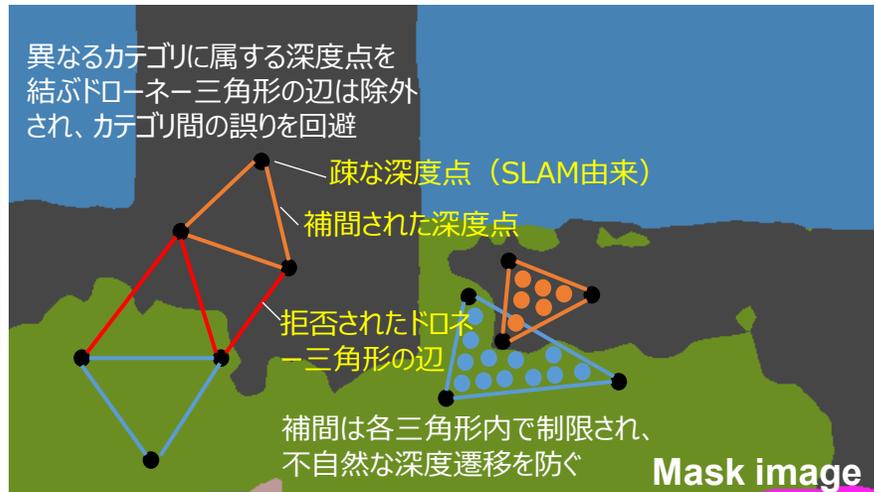
仮想世界



オクルージョン付き洪水MR



# 大規模都市環境におけるドローンMRのための オクルージョン処理手法





MR AI  
Metaverse

# Server-enabled Mixed Reality (2010–12, 2020–)

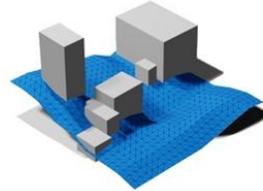
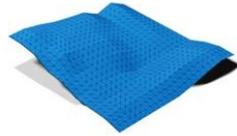
- 建造環境分野で使われる3Dモデルは、複雑でデータは大容量になりがち。
- MRクライアント（スマホ）には高性能なGPUなど高い演算能力が要求されてしまう。
- データ量を軽減するために3DモデルのLOD（Level of Detail: 詳細度）を自動制御する方法も検討されているが、適切なLOD処理をリアルタイムに実行することは難しい。
- 負荷の高い処理に対する他のアプローチとして、インターネットに接続された演算能力の高いサーバにオフロードできないか？

# 災害リスクを見える化：一般市民の理解を深めるMR

## デジタルツインとマルチクライアント対応によるオンサイト可視化

### 事前準備

0.2	0.1	0.0	0.3
0.5	0.3	0.0	0.2
0.7	0.7	0.2	0.1
1.0	0.8	0.4	0.1

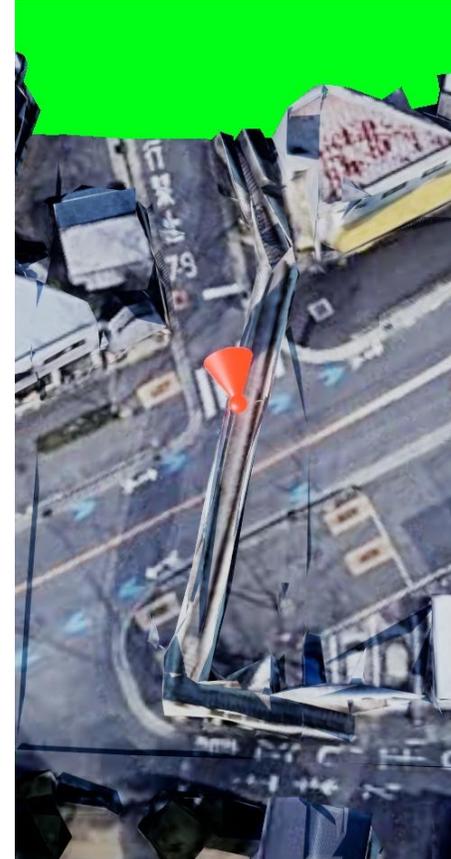
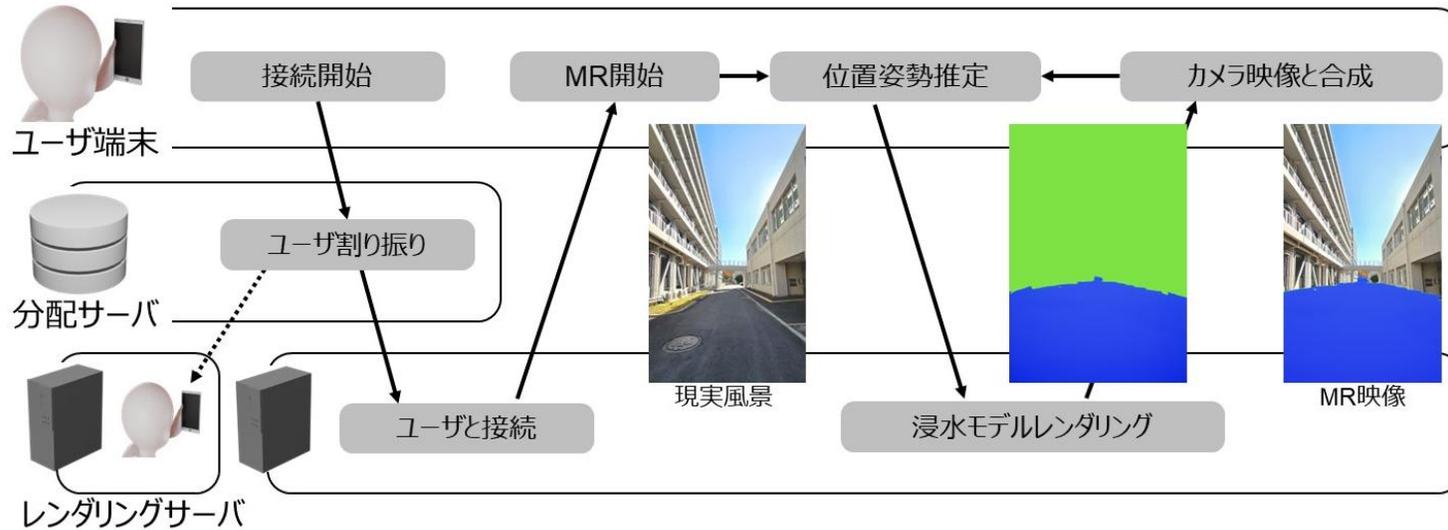


浸水シミュレーション

浸水面を3次元モデル化

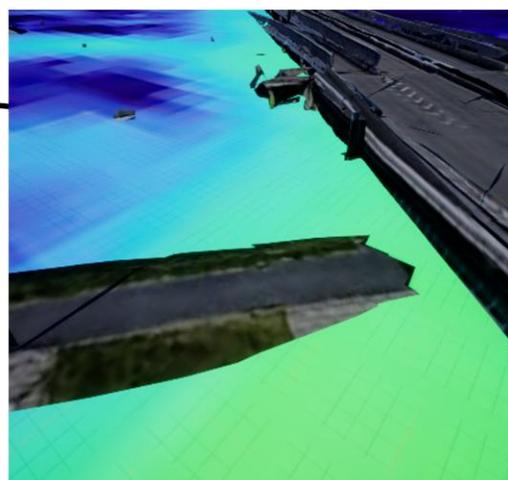
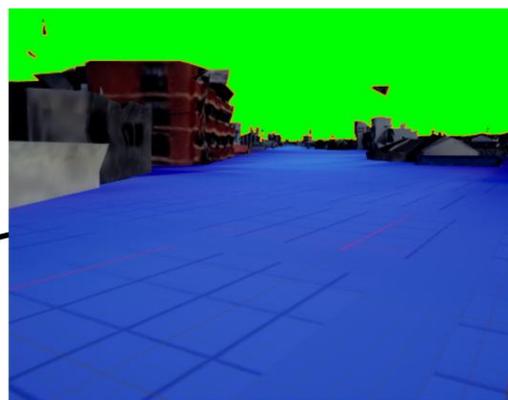
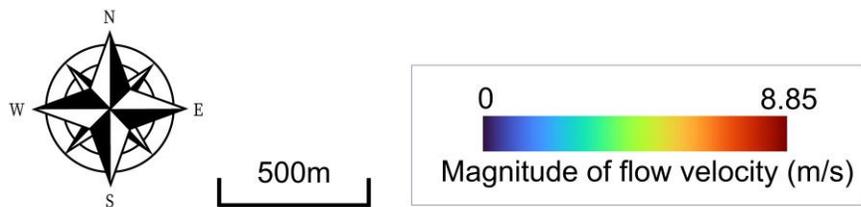
都市デジタルツインと合成

### MR表示



# 災害リスクを見える化：一般市民の理解を深めるMRシステム

デジタルツインとマルチクライアント対応によるオンサイト可視化



Flood flow velocity map

# マルチユーザー視点共有機能を含む Web型複合現実システムの開発



Camera image (現状)

Device A (チーフ)



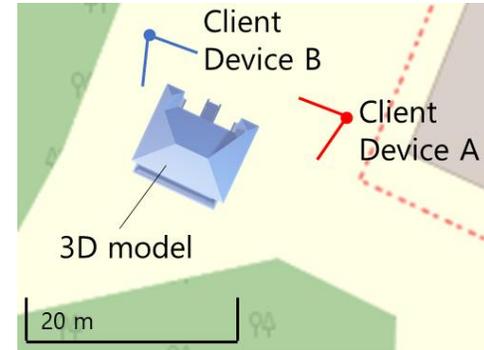
Server



Device B (担当者)



課題：何を見ながら、話をしているのかな・・・？



(サブ画面)担当者のMR



(メイン画面)チーフのMR

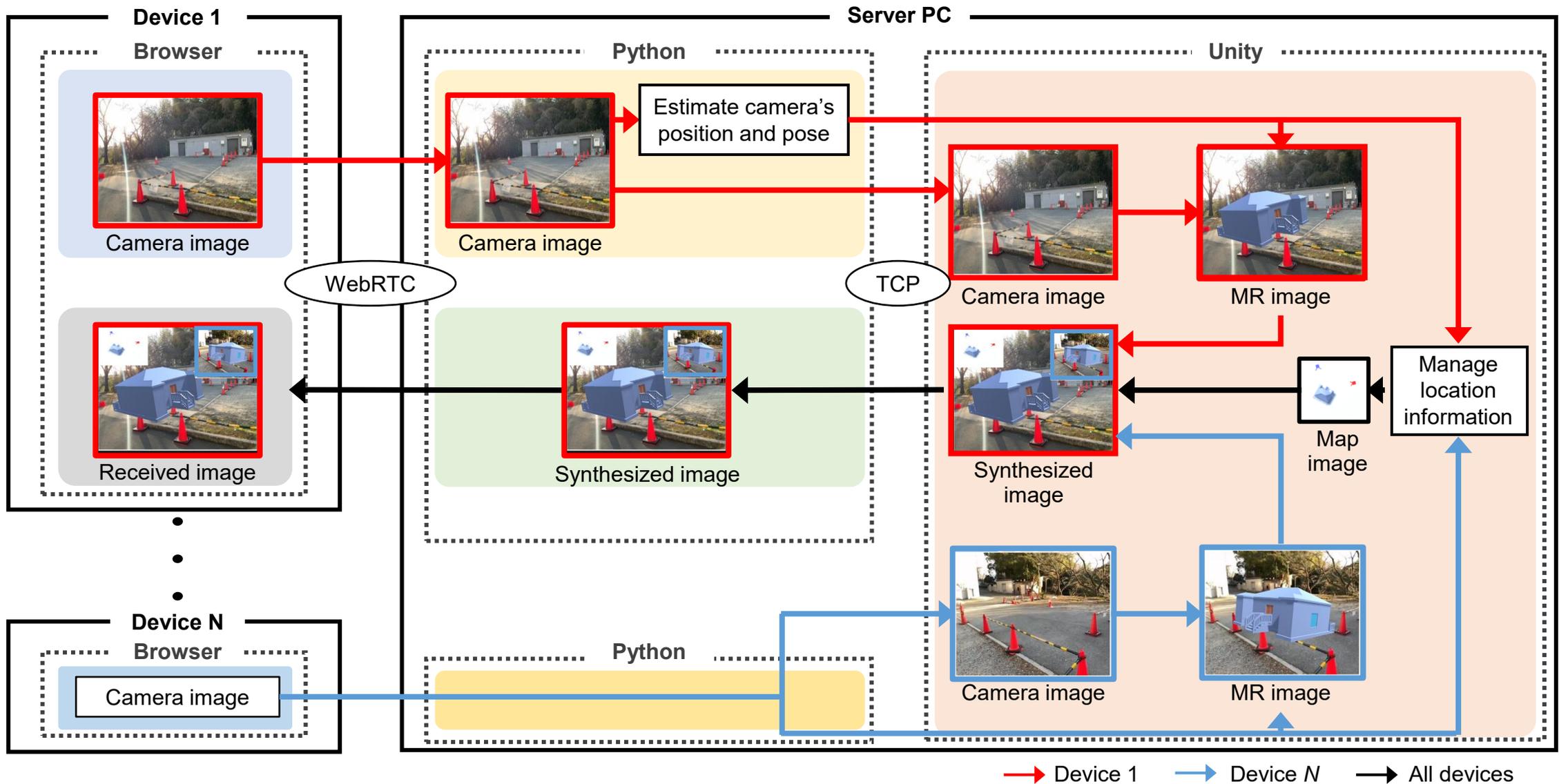


(メイン)担当者のMR画面

(サブ画面)チーフのMR

Synthesized image (MR)

# マルチユーザー視点共有機能を含む Web型複合現実システムの開発 -2





# AI 3D Modeling

# 設計モデル (To-be) と 現状モデル (As-is)

西淡町丸山公園環境設計 (兵庫県, 1994)



## 設計モデル (To-be)

3DCAD/CGソフトウェアで手入力

## 現状地形 (As-is)

国土数値情報 (国土地理院) の標高データ → 独自ソフトでメッシュ変換

## 現状建物 (As-is)

- 紙の地図をスキャン、ラスタライズ
- 建物平面形状 (フットプリント) を 3DCAD/CGソフトで手入力、階数・屋根形状ごとにレイヤー分類 (陸屋根、妻入り、入母屋など)
- 独自ソフトで3D形状化





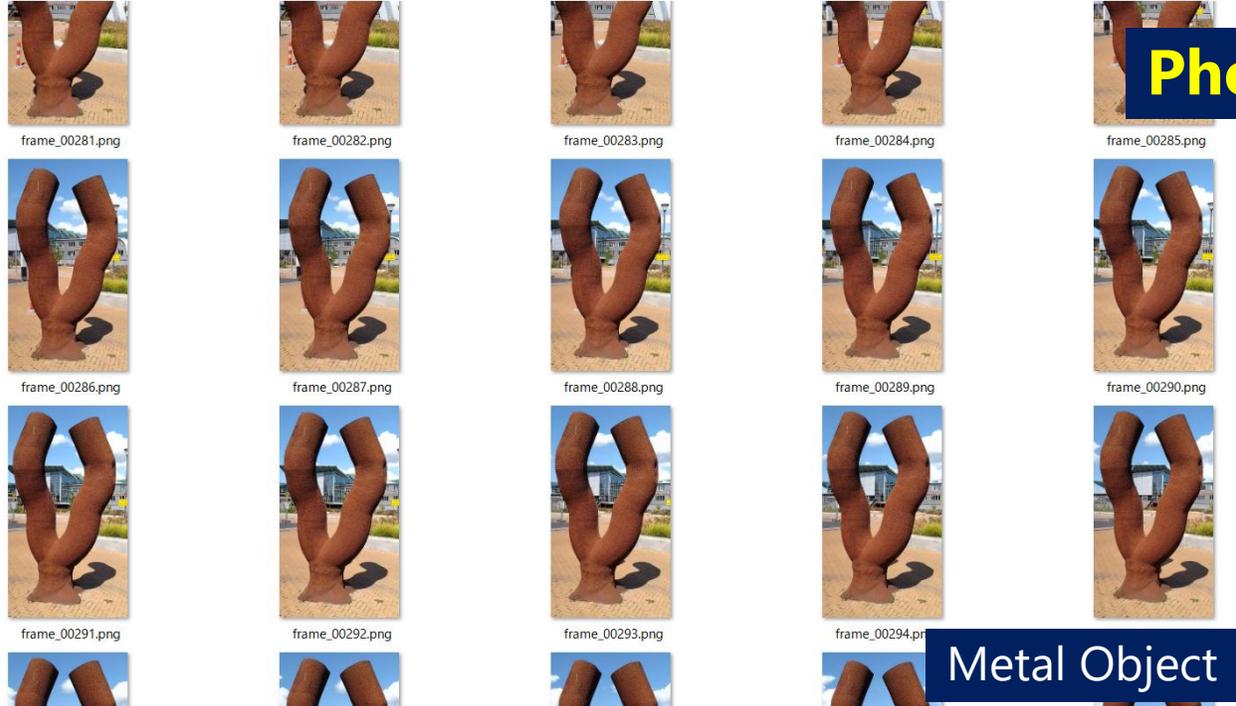
# 三次元再構成

3Dビジョンの一分野。特に視覚データやセンサーデータから実世界の物体や環境の3Dモデルを生成することに焦点

## 3D Gaussian splattering

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. 2023. 3D Gaussian Splattering for Real-Time Radiance Field Rendering. ACM Trans. Graph. 42, 4, Article 139 (August 2023), 14 pages. <https://doi.org/10.1145/3592433>

# Photos



frame\_00281.png

frame\_00282.png

frame\_00283.png

frame\_00284.png

frame\_00285.png

frame\_00286.png

frame\_00287.png

frame\_00288.png

frame\_00289.png

frame\_00290.png

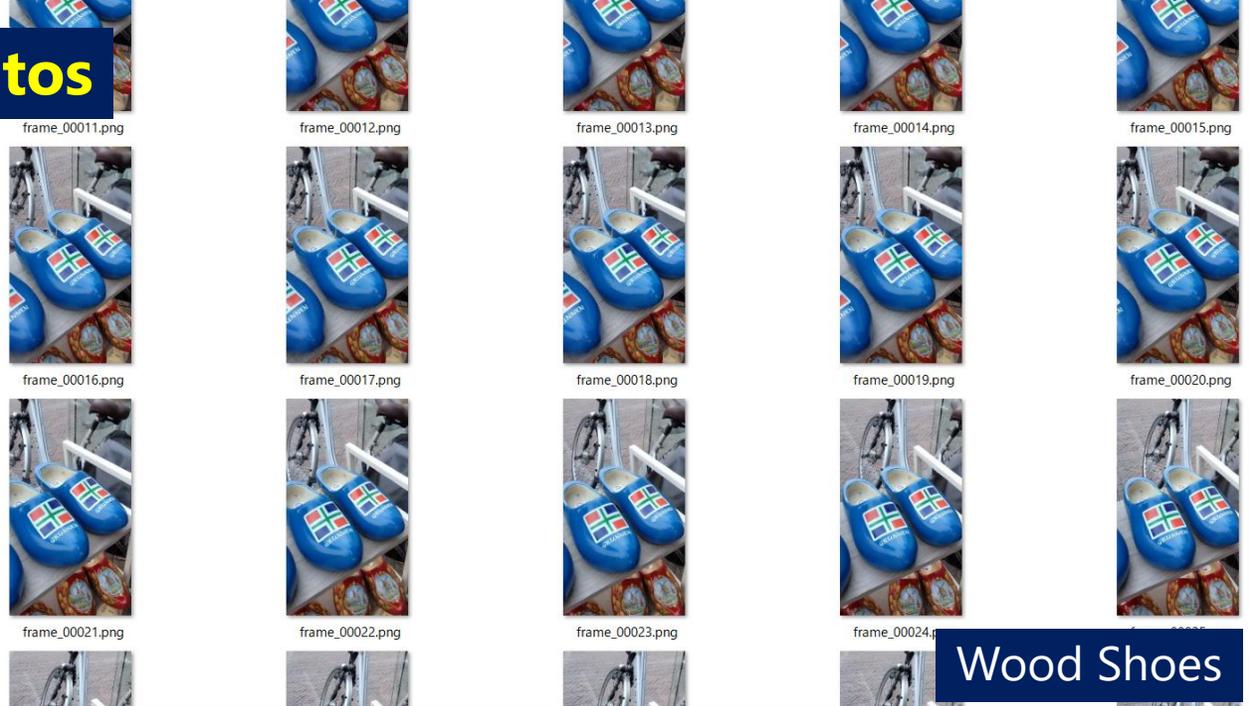
frame\_00291.png

frame\_00292.png

frame\_00293.png

frame\_00294.png

## Metal Object



frame\_00011.png

frame\_00012.png

frame\_00013.png

frame\_00014.png

frame\_00015.png

frame\_00016.png

frame\_00017.png

frame\_00018.png

frame\_00019.png

frame\_00020.png

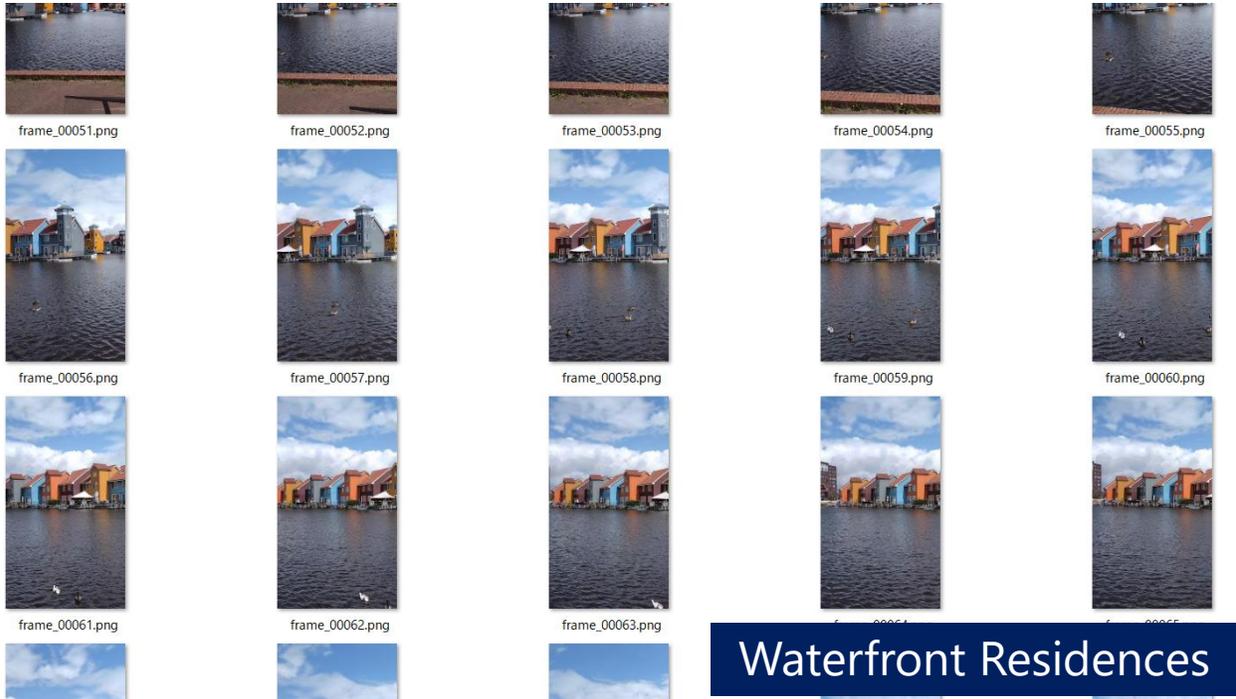
frame\_00021.png

frame\_00022.png

frame\_00023.png

frame\_00024.png

## Wood Shoes



frame\_00051.png

frame\_00052.png

frame\_00053.png

frame\_00054.png

frame\_00055.png

frame\_00056.png

frame\_00057.png

frame\_00058.png

frame\_00059.png

frame\_00060.png

frame\_00061.png

frame\_00062.png

frame\_00063.png

## Waterfront Residences



DSC\_4860.jpg

DSC\_4869.jpg

DSC\_4855.jpg

DSC\_4856.jpg

DSC\_4857.jpg

DSC\_4858.jpg

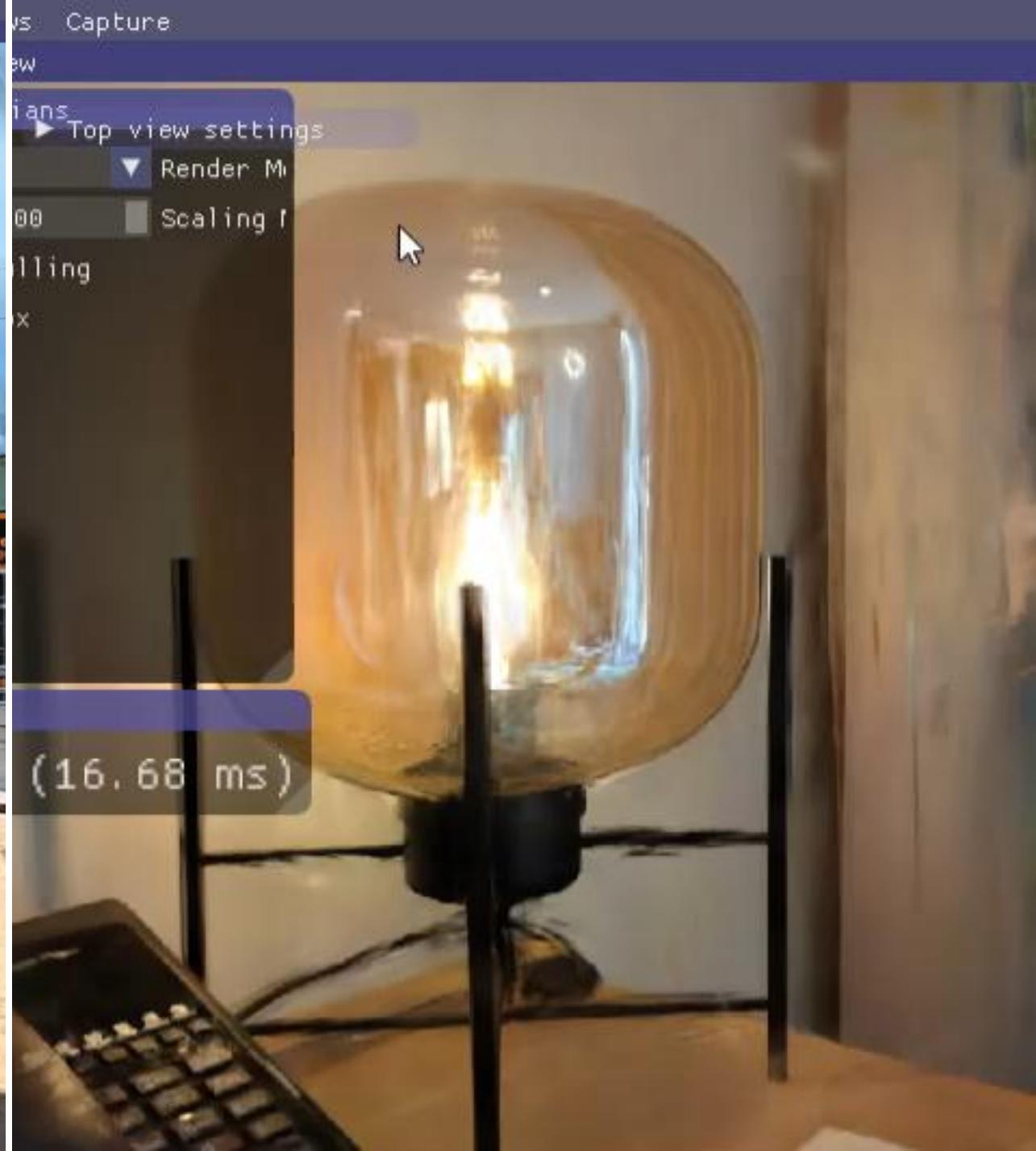
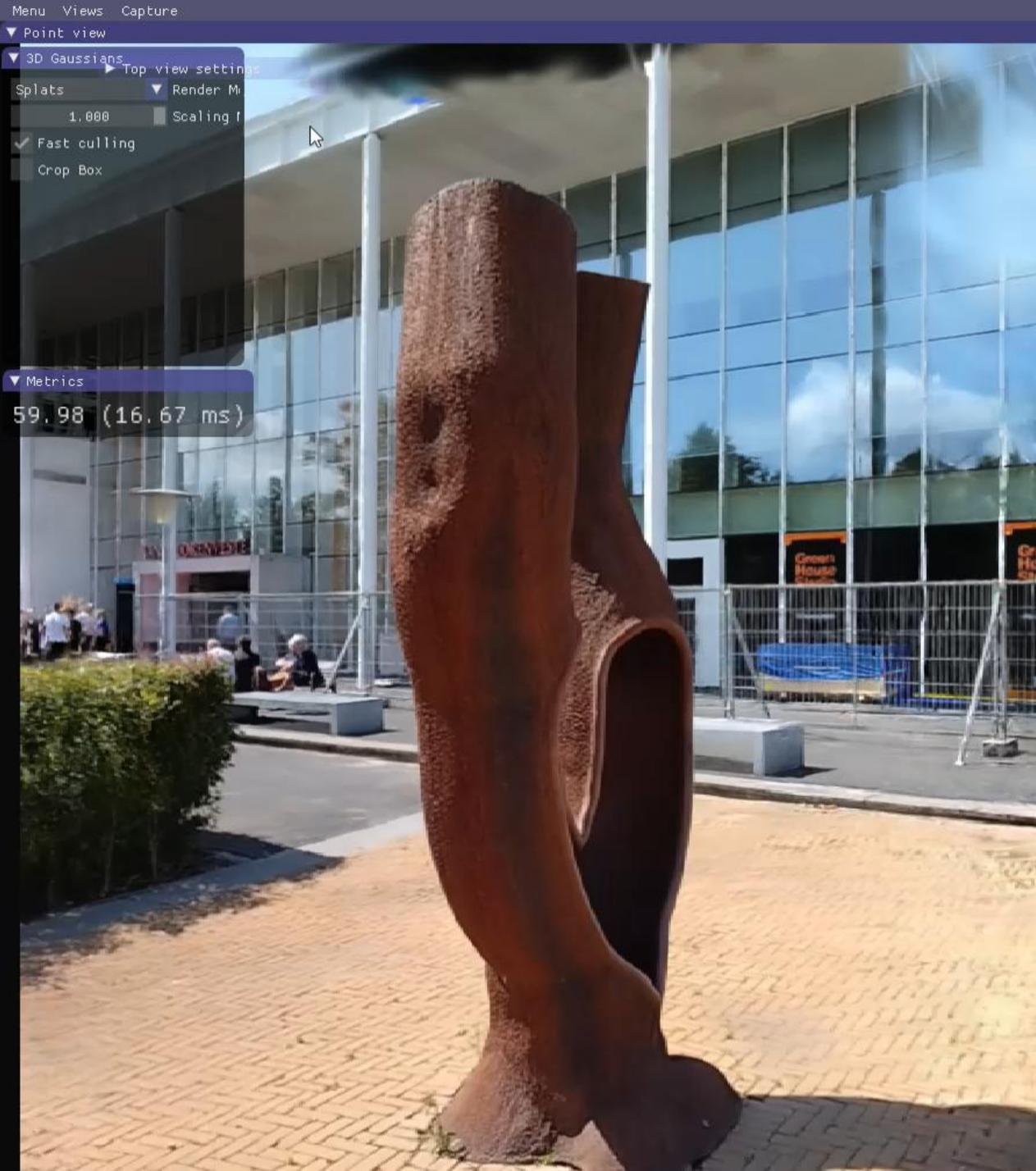
DSC\_4861.jpg

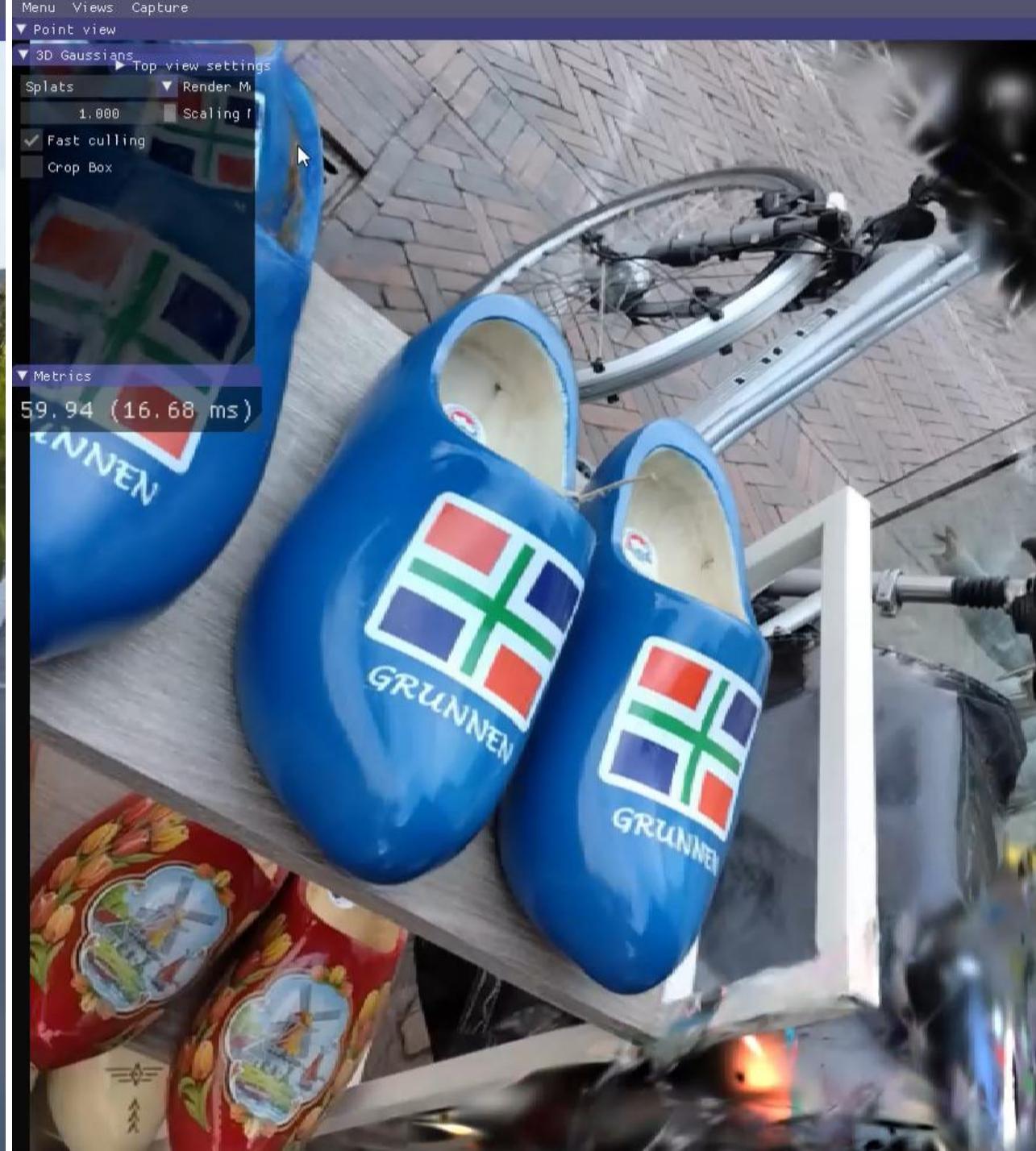
DSC\_4862.jpg

DSC\_4863.jpg

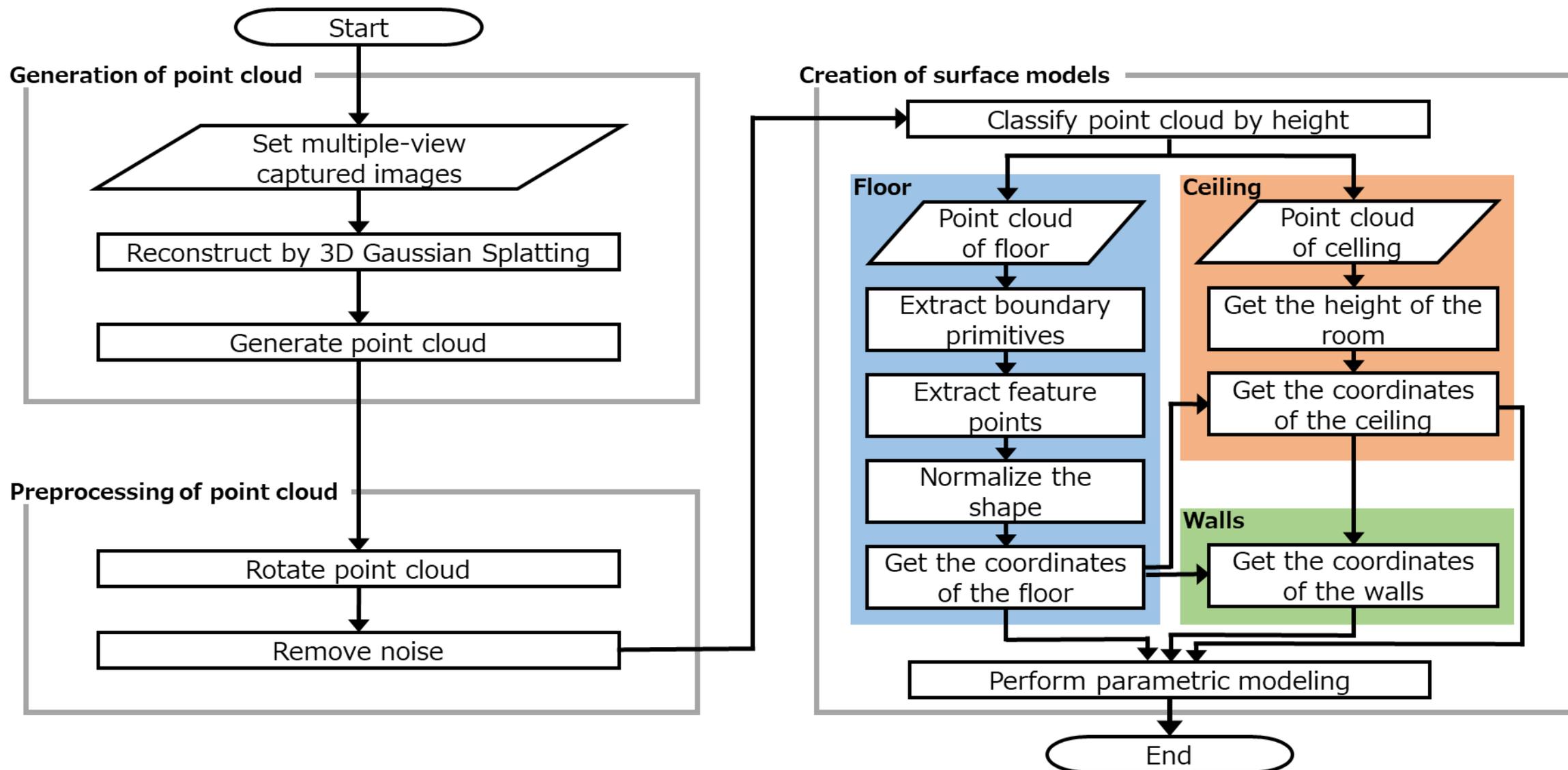
DSC\_4864.jpg

## Lump





# 3D Gaussian Splattingにより再構成した点群を用いた 屋内空間のサーフェスマデリング法



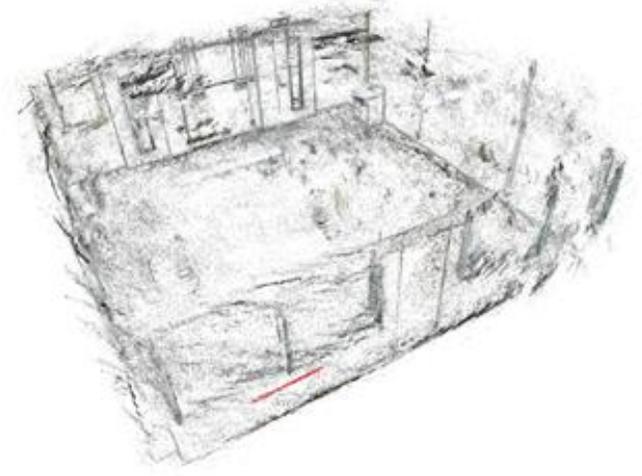
# 3D Gaussian Splattingにより再構成した点群を用いた 屋内空間のサーフェスマデリング法 -2



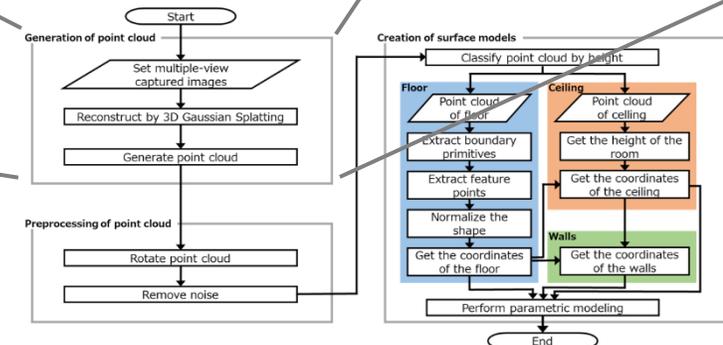
Multi-view captured images  
(153 images, 1600 px x 1198 px)



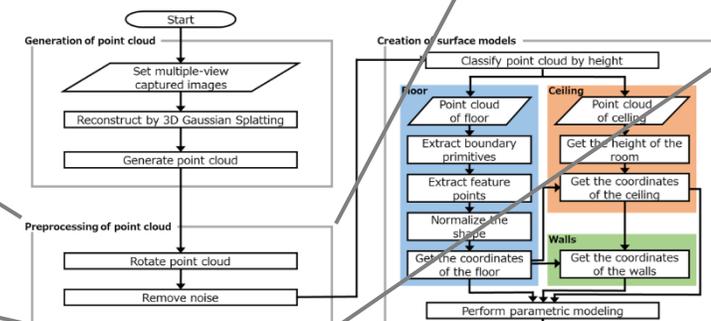
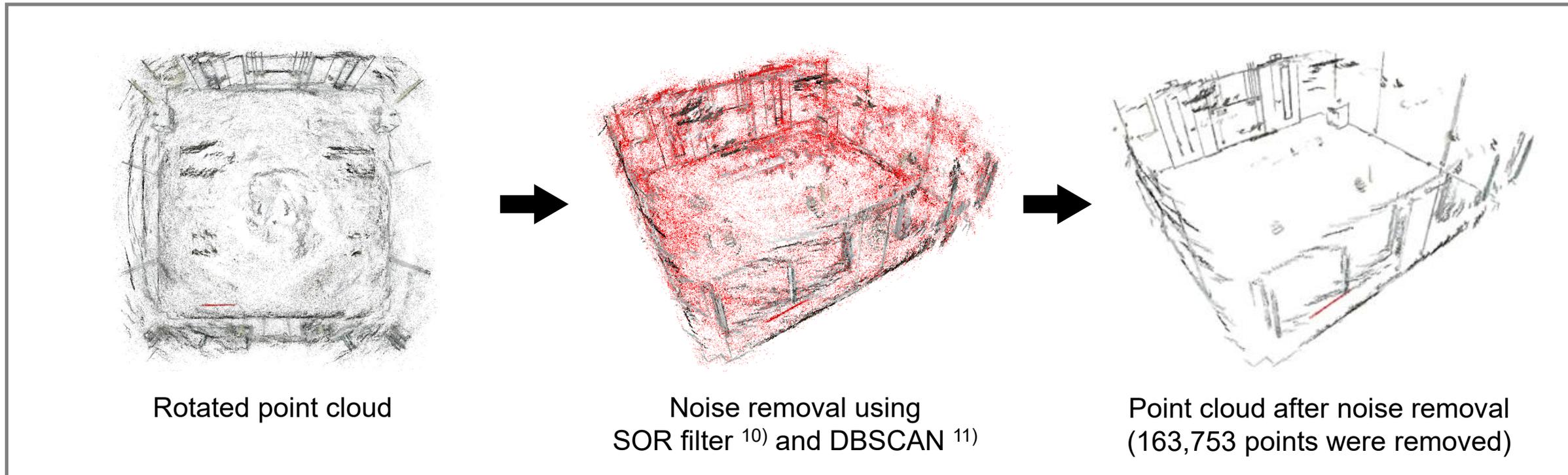
Reconstructed 3D scene  
by 3D Gaussian Splatting



Generated point cloud  
(Generation time: 3 min 39 sec)



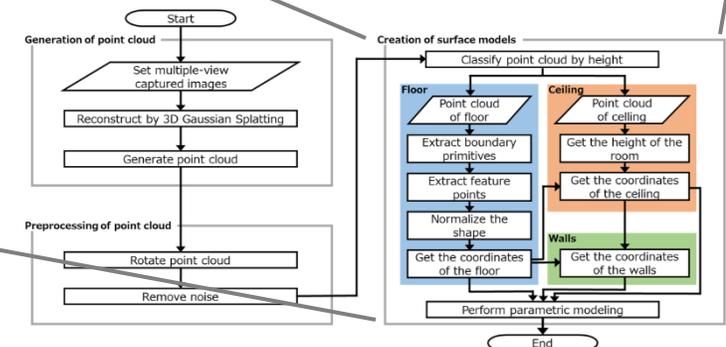
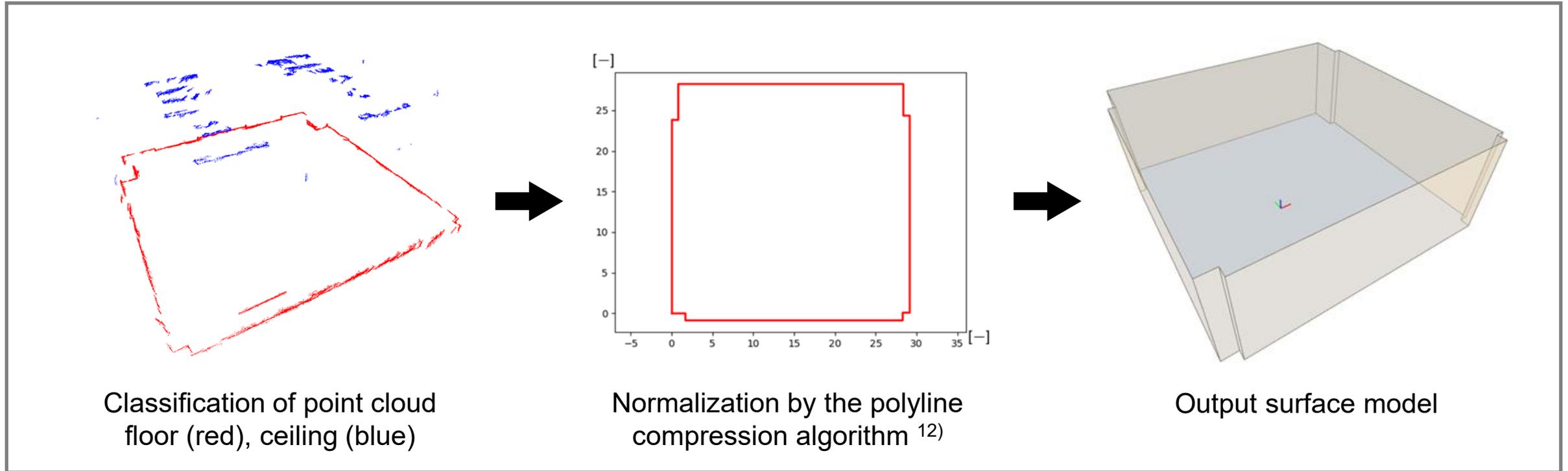
# 3D Gaussian Splattingにより再構成した点群を用いた 屋内空間のサーフェスマデリング法 -3



10) Rusu, R.B. (2010). Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. *Künstliche Intelligenz*, 24, 345–348. doi.org/10.1007/s13218-010-0059-6

11) Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Knowledge Discovery and Data Mining*.

# 3D Gaussian Splattingにより再構成した点群を用いた 屋内空間のサーフェスマデリング法 -4



12) Gribov, A. (2019). Optimal Compression of a Polyline While Aligning to Preferred Directions. 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), 1, 98-102.

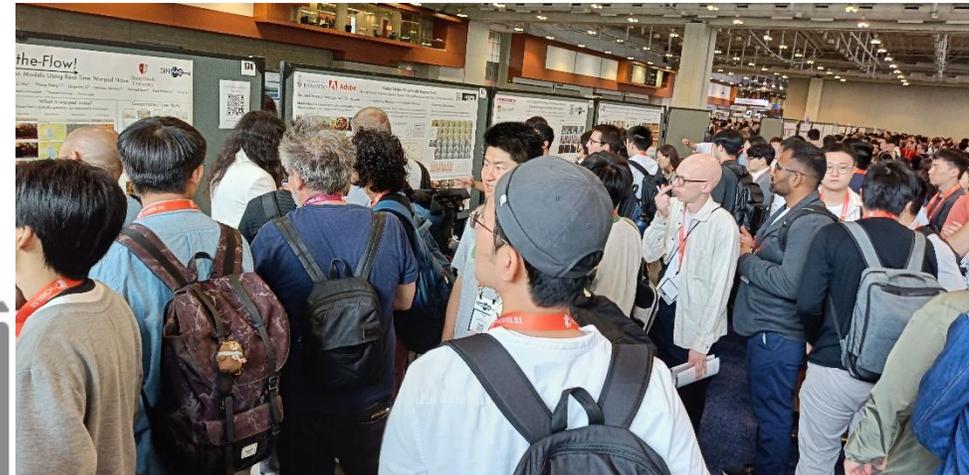
# CVPR 2025, Jun. 11-14

The highest int'l conference on image-based AI:

- The h-5 index (a metrix of the  $N$  of papers cited  $N$  times or more in the past five years) is **440**, second only to Nature magazine.

2025

- Paper submission: 13,008
- Accepted: 2,872 (Acceptance rate: 22%)
- 12,000 participants from 76 countries/regions
- Poster sessions: in the morning and afternoon, with a total of 1,000 posters per day, 500 posters per 2 hours!



Google Scholar

Top publications

Categories English

Publication	h5-index	h5-median
1. Nature	488	745
2. IEEE/CVF Conference on Computer Vision and Pattern Recognition	440	689
3. The New England Journal of Medicine	434	897
4. Science	409	633
5. Nature Communications	375	492
6. The Lancet	368	678

# VGGT: Visual Geometry Grounded Transformer

## Object

<https://huggingface.co/spaces/facebook/vggt>



### 1. Input: Upload a video or image.

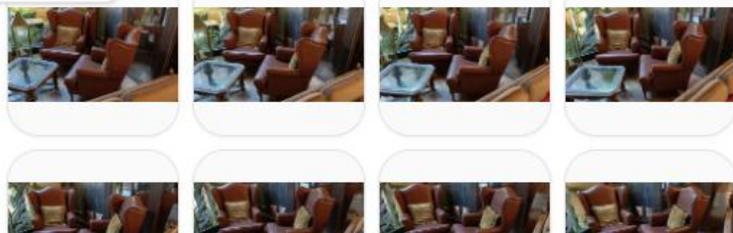
Upload Images

1920 x 1080, 20 seconds  
Shot: July 16 by TomoFuku

ファイル  
ここにドロップ  
または  
クリックしてアップロード

### 2. Generate 20 frames

Preview



3D Reconstruction (Point Cloud and Camera Poses)

Upload complete. Click 'Reconstruct' to begin 3D processing.

3D Model



Reconstruct **3. Reconstruct**

Clear

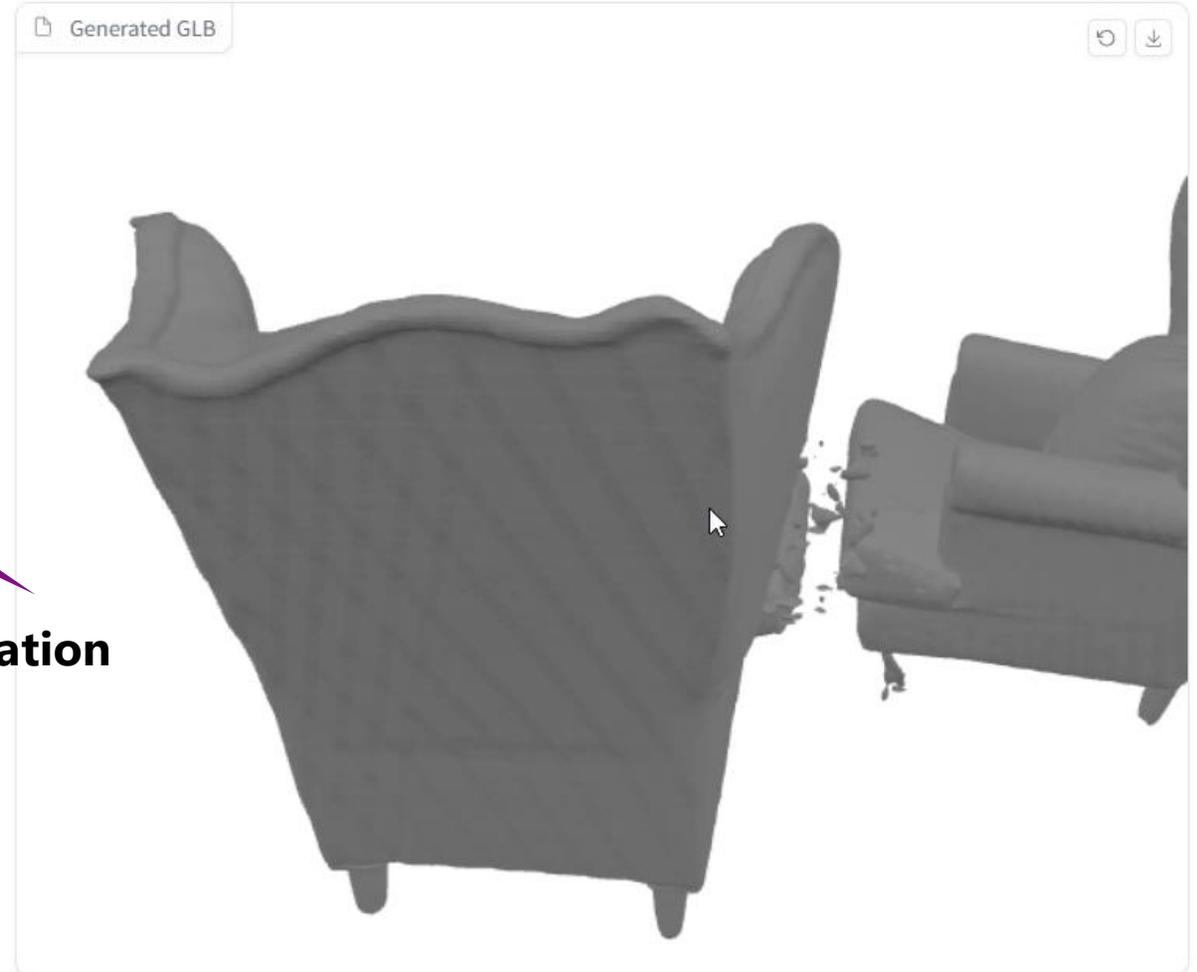
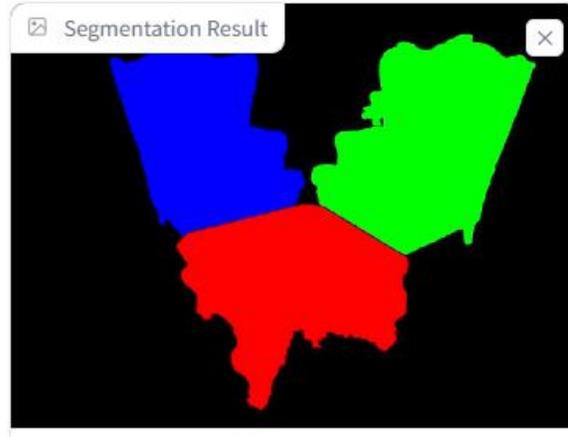
"Our model itself usually only needs less than 1 second to reconstruct a scene. However, visualizing 3D points may take tens of seconds due to third-party rendering, which are independent of VGGT's processing time."

Depthmap and Camera Branch

Pointmap Branch

# MIDI: Multi-Instance Diffusion for Single Image to 3D Scene Generation

<https://huggingface.co/spaces/VAST-AI/MIDI-3D>



1. **Input: Upload an image.**  
**Draw bounding boxes**

Segmentation Settings 4624 x 3468  
Shot: July 16 by TomoFuku  
 Polygon Refinement

2. **Run Segmentation** Run Segmentation

Generation Settings

Do image padding

Seed

2011819116

3. **Generation**

Download GLB

# MoGe-2: Accurate Monocular Geometry with Metric Scale and Sharp Details

<https://huggingface.co/spaces/Ruicheng/MoGe-2>



Input Image



## 1. Input: Upload an image.

Settings

4624 x 3468

Maximum Image Size

Shot: July 16 by TomoFuku

800

Inference Resolution Level

High

3D View Depth Normal Measure Download

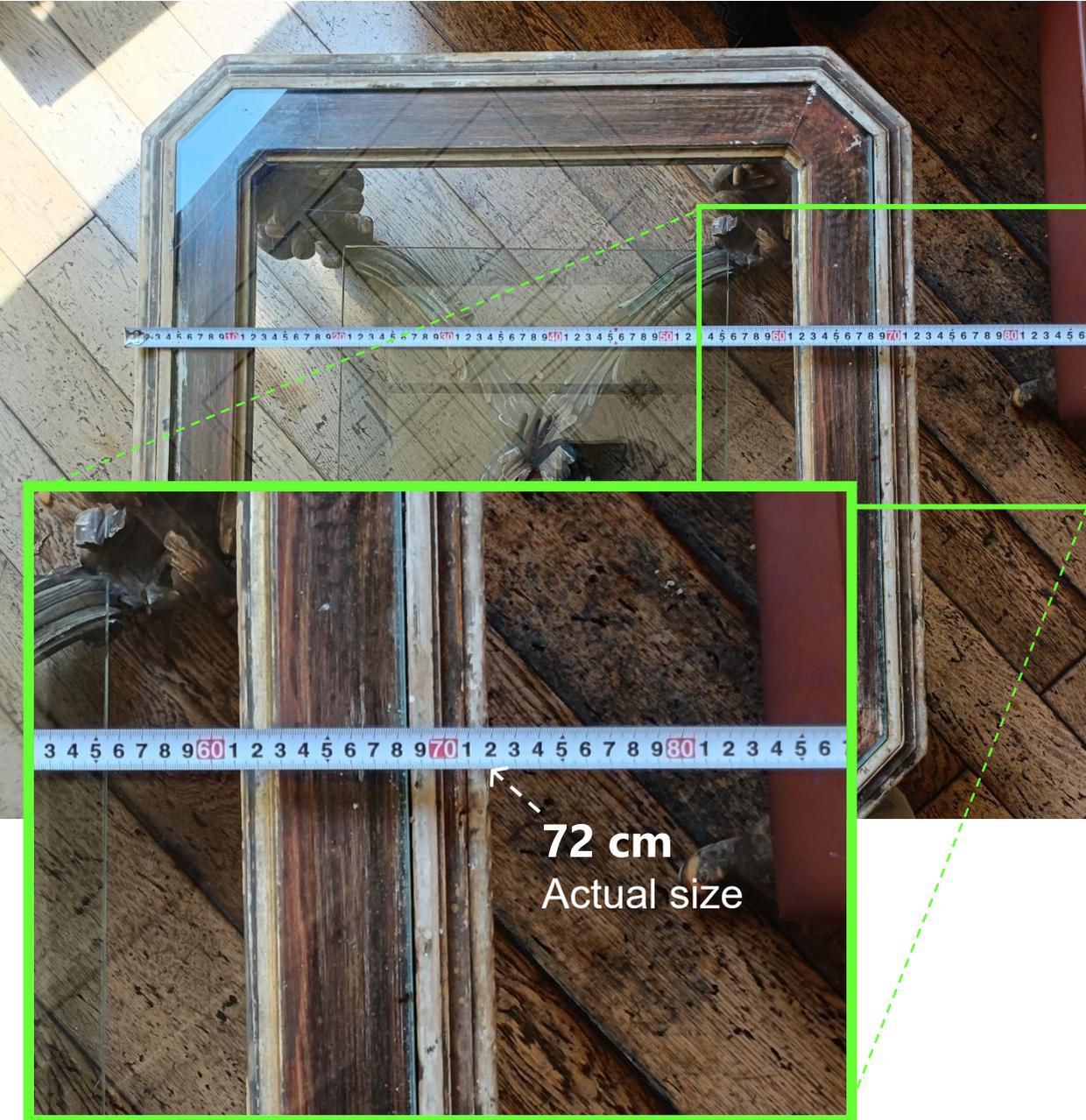
**Note:** Inference has been completed. It may take a few seconds to download the 3D model.

3D Point Map



# MoGe-2: Accurate Monocular Geometry with Metric Scale and Sharp Details

<https://huggingface.co/spaces/Ruicheng/MoGe-2>



- P1 depth: 1.49m.
- P2 depth: 1.75m.
- Distance: 0.73m

**73 cm**  
Measured with MoGe-2

# MoGe-2: Accurate Monocular Geometry with Metric Scale and Sharp Details

<https://huggingface.co/spaces/Ruicheng/MoGe-2>



St. Peter's Basilica seen through a keyhole at the Villa Malta  
マルタ騎士団の鍵穴

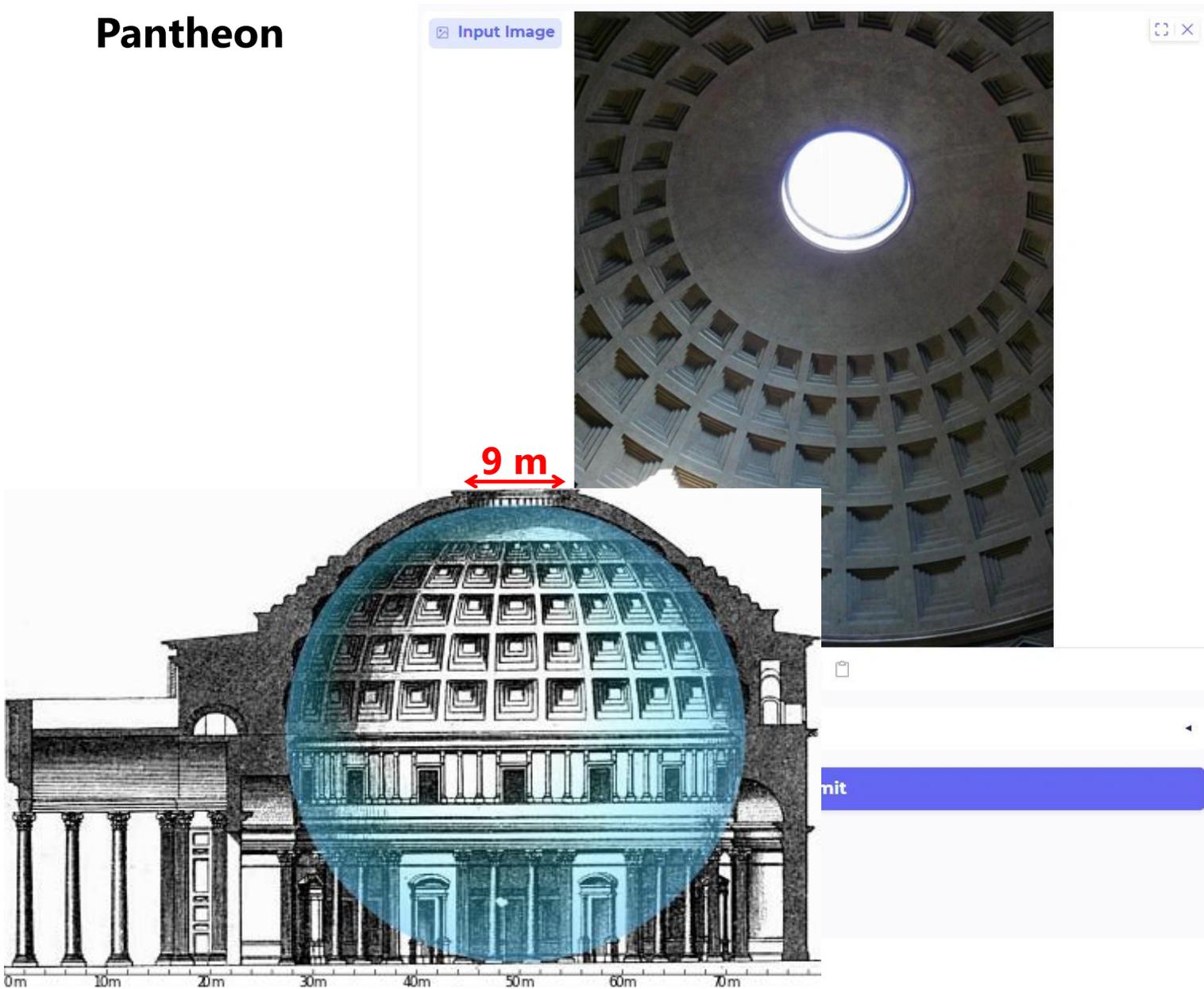


**Dome diameter**  
**41 m (Actual size)**  
**16.34 m (Measured with MoGe-2)**

# MoGe-2: Accurate Monocular Geometry with Metric Scale and Sharp Details

<https://huggingface.co/spaces/Ruicheng/MoGe-2>

## Pantheon

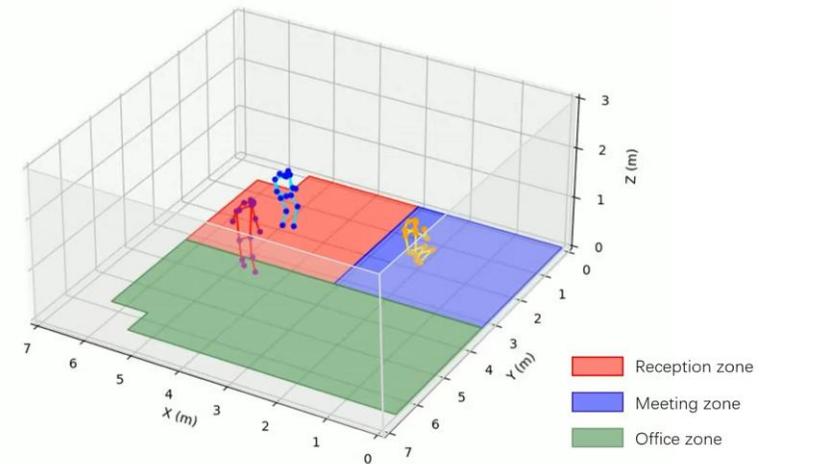
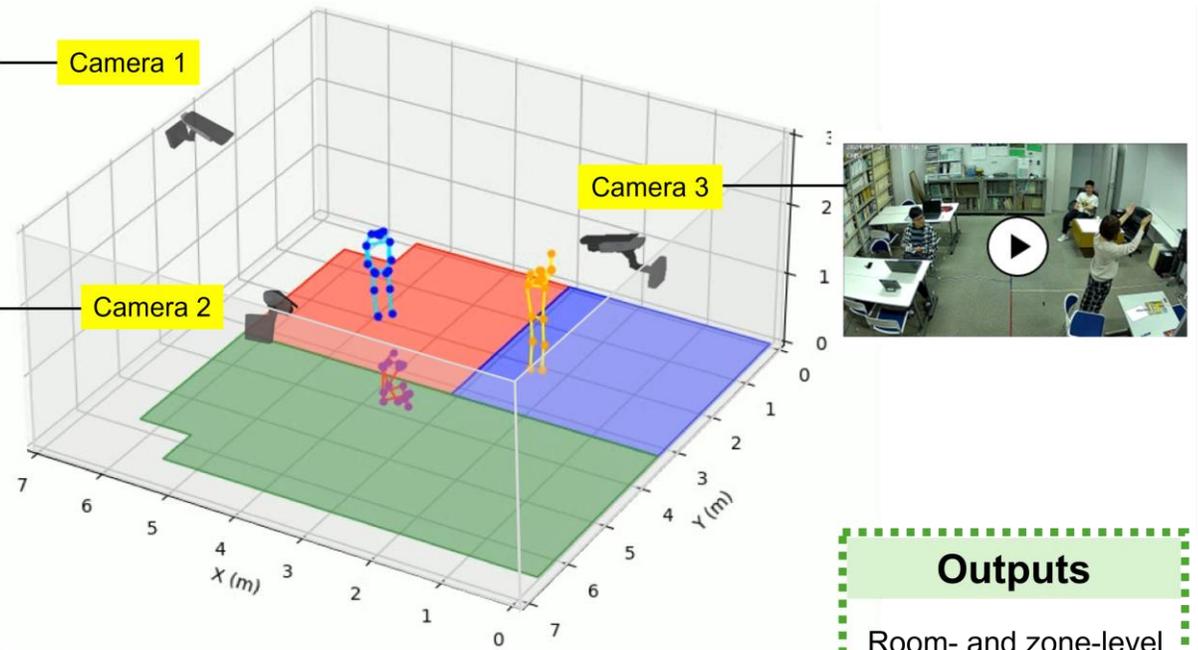


空間利用、AIに聞け！？

# オープンオフィス空間におけるゾーンごとの占有状況測定システムを開発

—3次元姿勢推定を活用し、低コストで高信頼性データを提供—

- ✓ 開放型オフィス内の細かいエリアごとの使用状況をリアルタイムで精緻・定量的に把握するフレームワークを構築。
- ✓ 細かいエリアごとの占有状況を的確に把握する有効な手法は未確立だったが、複数視点からのビデオカメラ画像について人体関節位置による3次元姿勢推定（3DPE）を行うことで、ゾーンごとの占有状態を判定することが可能に。
- ✓ オープンオフィス環境における信頼性の高い空間利用データの提供により、レイアウトの最適化、資源管理、エネルギー制御などオフィス環境における持続可能な設計・運用の高度化支援の発展に期待。

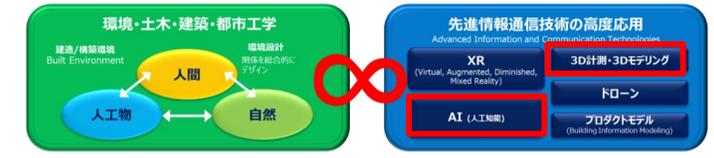
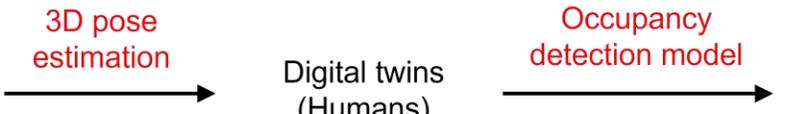


**Inputs**

- Interior space
- Muti-view videos
- A floor plan

**Outputs**

- Room- and zone-level
- Occupancy status
- Occupancy metrics



Chen, S., Fukuda, T., Yabuki, N. (2025). Development of an occupancy measurement system for micro-zones within open office spaces based on multi-view multi-person 3D pose estimation, Journal of Building Engineering, Volume 111, 113037, <https://doi.org/10.1016/j.job.2025.113037>



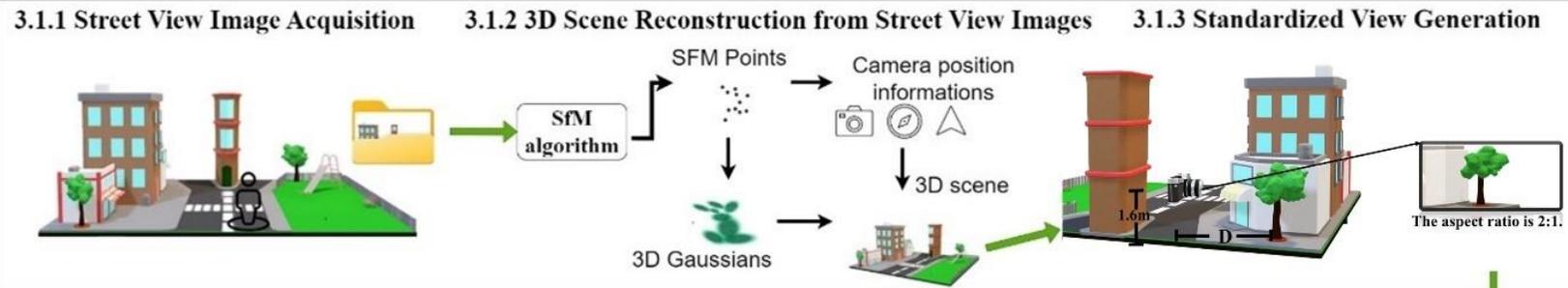
A I

# 街の緑の評価を、「量」から「質」へ 深層学習による都市緑地の高精度評価システムを開発

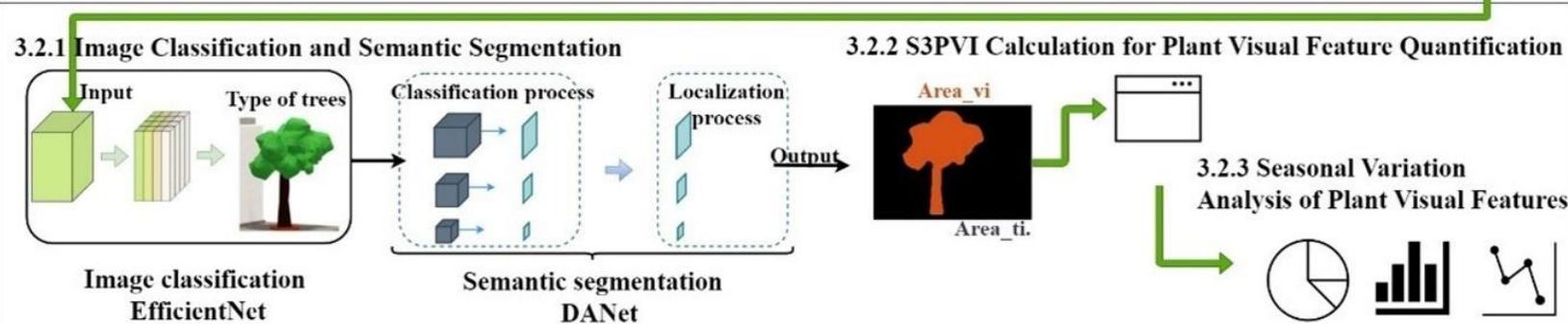
—多時相解析フレームワークでスマートに季節感あふれる緑地計画支援—

- ✓ ストリートビュー画像を活用し、都市の緑地における植生を、複数の時点（多時相）で可視化・解析するフレームワークを開発。植物種ごとの季節変化を定量的に捉える革新的な技術アプローチを実現。
- ✓ 新たに提案した「季節種別植物景観指数（S3PVI）」により、51種の都市植物データセットで平均82.17%の高精度識別を達成。従来の緑視率では困難だった植物種別の詳細評価を可能に。
- ✓ 深層学習と3D再構成技術（SfMおよび3Dガウススプラッシング）を統合。標準視点の自動生成により、解析精度とデータの一貫性を大幅に向上。
- ✓ 大阪府吹田市での実証実験により技術の有効性を確認。今後の都市緑地評価・計画支援に資する基盤技術としての応用が期待される。

## 3.1 Data Collection and Preprocessing



## 3.2 Plant Visual Feature Evaluation Based on S3PVI



Hu, A., Yabuki, N., Fukuda, T. (2025). A Multi-temporal Framework for Urban Green Space Vegetation Visualization and Analysis Using Deep Learning and 3D Reconstruction, *Landscape Ecology*, 40, 125. <https://doi.org/10.1007/s10980-025-02090-4>

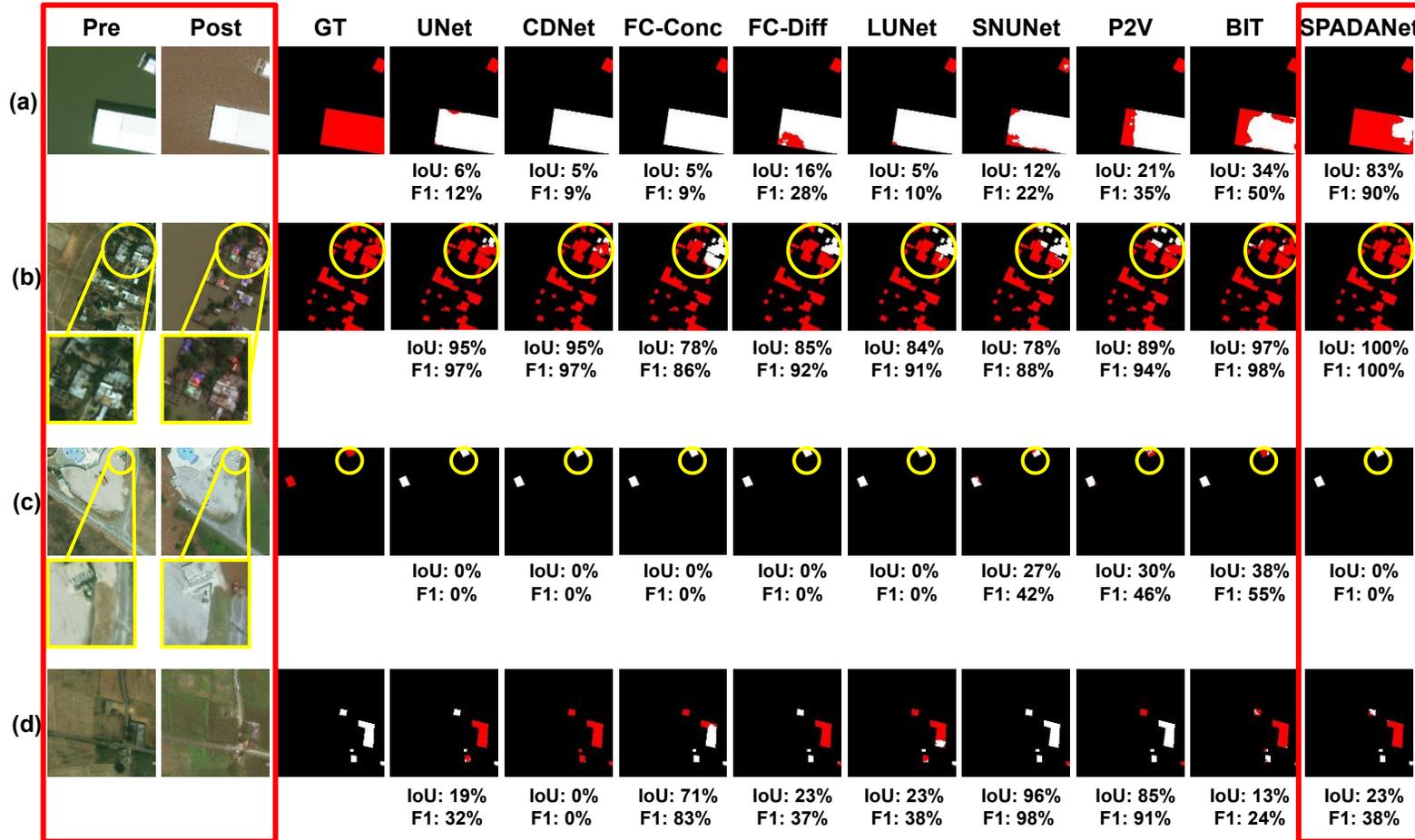
「見えない建物被害」も見抜く！？

# 水害後建物被害評価AIの新たなベンチマークを開発

—見落としを防ぐ「再現率」重視のアプローチで人命救助に貢献—



- ✓ 衛星画像を用いた水害後の建物被害評価（Flood-BDA: Building Damage Assessment）に特化した深層学習モデルの性能を評価するための、世界初の体系的なベンチマークを確立
- ✓ 災害発生直後に大量のラベル付けされた訓練データを入手することは難しいが、新たに開発した半教師あり学習（SSL）の手法は、わずか10%のラベル付きデータで完全教師あり学習の約74%の性能を達成し、水害シナリオにおける「画像レベル一貫性正則化（Consistency Regularization）」手法の有効性を実証
- ✓ 提案した軽量の深層学習モデル「Simple Prior Attention Disaster Assessment Net (SPADANet)」は、「被災建物の見逃し」という致命的なエラーを大幅に削減し、再現率をベースラインモデルと比較して約9%向上させた
- ✓ 将来の災害対応AIに不可欠な設計原理を提示し、人命救助活動のための、より迅速かつ効率的な深層学習モデルの開発に貢献



- Presence
- Interaction
- Autonomy

Visualization

Design study  
Consensus building



Output MR



Live video

Sparse point cloud

Semantic segmentation

Depth map

### 大規模都市環境におけるドローンMRのための オクルージョン処理手法

Modeling

### 3Dガウスプラットフォームを用いた 屋内サーフェスモデリング法

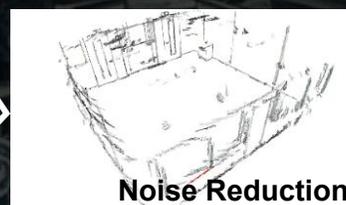


Photos from multiple viewpoints

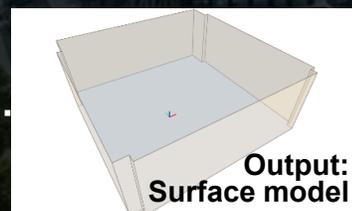
Design acts



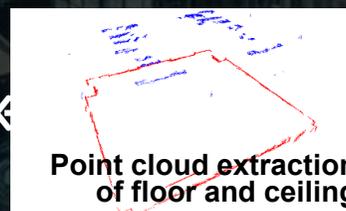
Point cloud by 3DGS



Noise Reduction

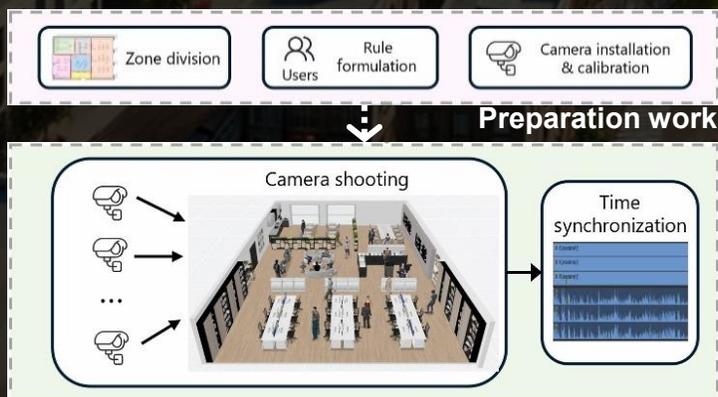


Output:  
Surface model

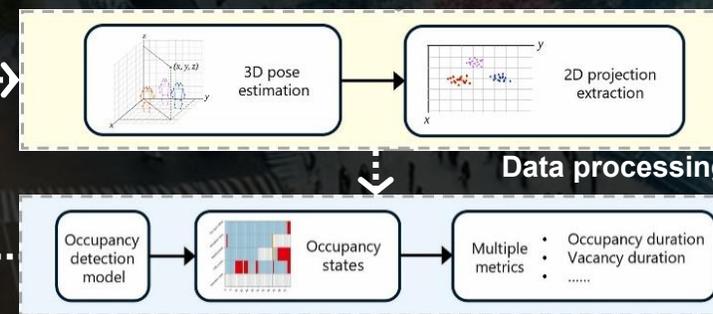


Point cloud extraction  
of floor and ceiling

Monitoring  
Analyzing



Layout optimization by zone function



### 多視点多人数3D姿勢推定に基づく オープンオフィス内マイクロゾーンの 占有率計測システムの開発

# 平時・災害両立 フェーズフリー

Compatibility in normal times and disasters



本講演に関わった全ての皆様に感謝申し上げます。

Thanks to all colleagues who participated projects in this presentation.