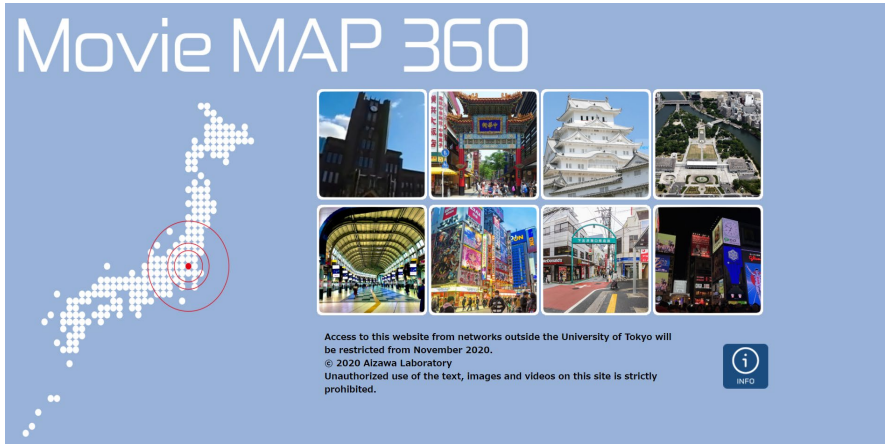
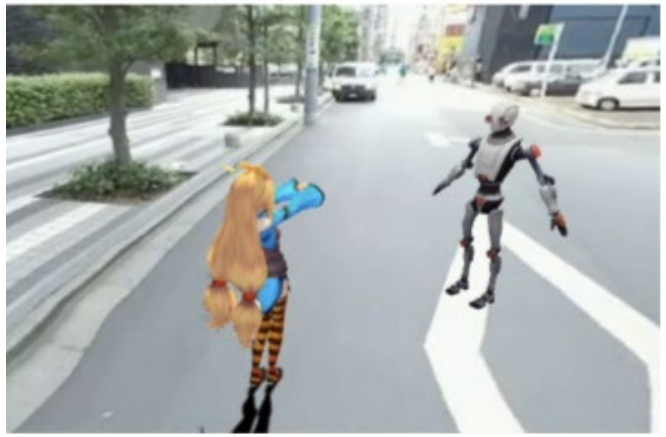




Building a Realistic Virtual World from 360° Videos for Large Scale Urban Exploration

Kiyoharu Aizawa
Dept. of EEIC, VR Center
The University of Tokyo,

VCIP 2024

Content

- Real World Metaverse
- Virtual World : CG vs Video
- Photo-realistic Representations of Urban Areas

- Movie Map
- 360RVW
- 360xCityGML RVW

- Conclusion

Real World Metaverse

Real World Metaverse

- John Hanke, Niantic CEO

August 10, 2021

The Metaverse is a Dystopian Nightmare. Let's Build a Better Reality.

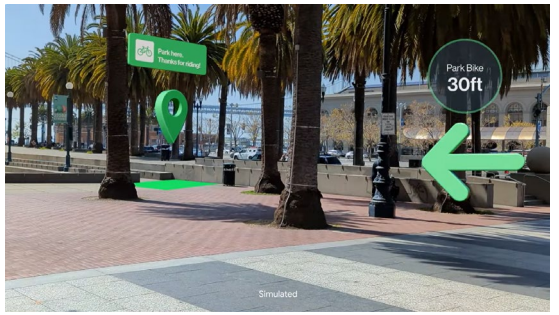
We believe we can use technology to lean into the 'reality' of augmented reality — encouraging everyone, ourselves included, to stand up, walk outside, and connect with people and the world around us.

General Issues about connection btw the real and metaverse.

Approaches to the integration of real and VR world

- Augmented Reality（AR）
  - Pokemon Go, Google AR
  - Apple Vision Pro, Microsoft HoloLens
- Virtualized Reality
  - Virtual towns, ex. Virtual Shibuya etc.
  - Video, Point cloud, Volumetric video
- Generative AI
  - text-to-3D, image-to-3D

Pokemon GO



Google

Virtual Shibuya, cluster

Make-It-3D, ICCV23

Choices of technology for incorporating the real world into VR

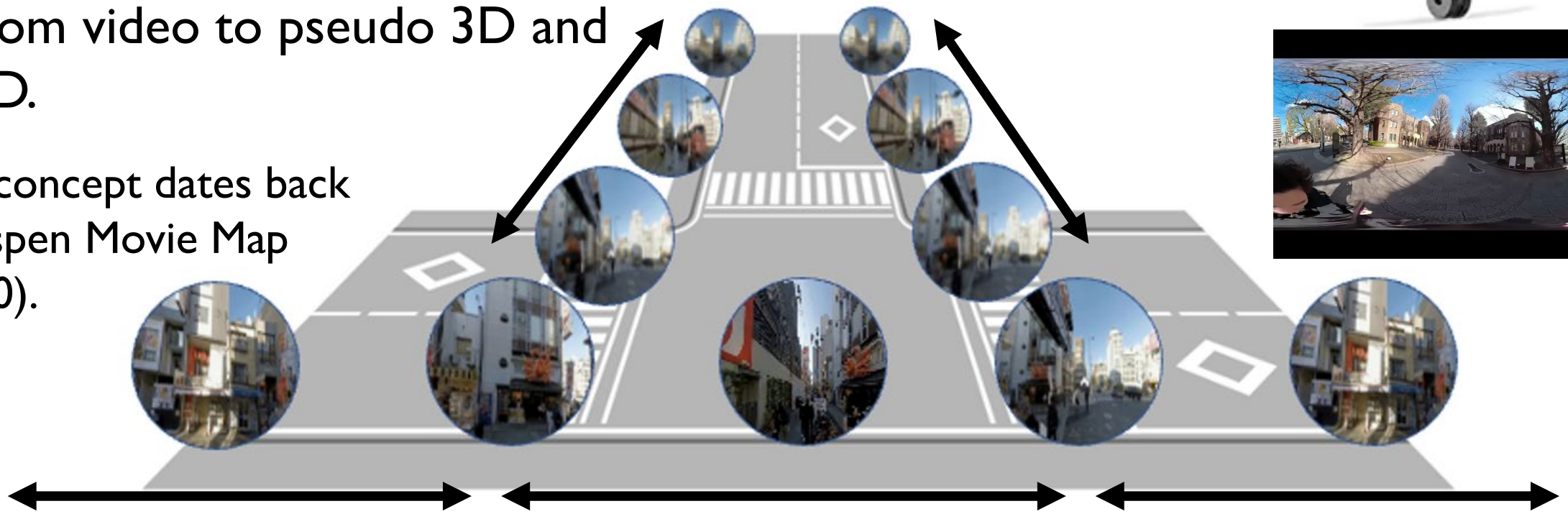
Real 360° Video

~ Virtual Exploration of Urban Areas ~

360°Video for Creating Virtual Urban World

- Movie Version of Google Street View (GSV).
- Connecting video segments builds Virtual Urban World.
- The representation can evolve from video to pseudo 3D and 3D.

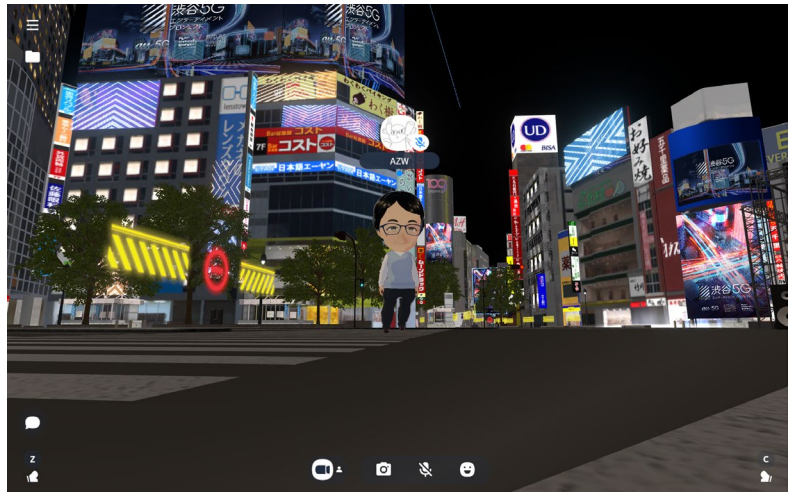
The concept dates back to Aspen Movie Map (1980).



Demo : 360RVW for PC

Demo

CG vs Video : Virtual Worlds for Real Cities



CG ⟷ Video

Graphic based Virtual World
Ex. Virtual Shibuya

360RVW: Fusing 360 ° Videos and Virtual World

Modeling (Space)	Heavy	Light	Modeling Free
Modeling (Moving Object)	Difficult	Easy	
Computation	Heavy	Light	Easy to Scale up
Photo-Reality	Low	High	
Interactions	High	Low	

@Univ of Tokyo

The Importance of Real Video

Authenticity: Real video ensures the content feels real and genuine.

The advantages of using real 360° video

- Unlike Google Street View or CG models, video captures the movement of scenes and conveys the atmosphere of the place.
- No need for time-consuming modeling -- just take the video.
- Streaming video removes the need to download large files.
- The amount of user-side computation is low, making it suitable for scaling up.

video segment data made from 360° videos in urban areas

Video data is shared

Movie Map

- Interactive video that moves around the urban areas
- Accessible from a web browser
- In operation by **moviemap.jp**

360 RVW

- Pseudo 3D representation
- Introducing avatars and objects
- Accessible via a browser using WebGL
- Extended to a smartphone app

360xCityGML RVW

- Use of global city models (Plateau, CityGML)
- Dynamic video mapping
- 3D visualization such as flood etc.

Photo-realistic Representations of Urban Areas

Photo-realistic Representation of Urban Areas

Static Image Repr.

- GSV (2007～)
- 3D + GSV texture
  - Geollery (2019)
  - Avatar360 (2024)
  - FLY (2024)

Video Repr.

- AspenMovieMap (1980)
- Movie Map (2020)
- 360RVW (2023)
- 360RVW+CityGML (2024)

NeRF, Gaussian Splatting

Novel view generation from Multiview Images

Drive/Walk through
- Block-NeRF (2022)
- SUDS (2023)
- Entity-NeRF(2024)

Fly through
- CityNeRF (2022)
- Mega-NeRF (2022)
- CityGaussian(2024)

Aspen Movie Map (1980)

- Interactive drive though video of Aspen city (1980), long before GSV.
- Analog technology using a panoramic film camera and analog optical disks.
- Every street movie and turns at every intersection captured.
- Due to the tremendous effort required, there was no follow-up after Aspen.

Aspen Interactive Movie Map
https://www.youtube.com/watch?v=2Ytd12d6qNw (SIGGRAPH80)

Google Street View (GSV) (2007～)

- Sparsely located 360° still images
- Global coverage
- Mapillary (2013~) based on crowdsourcing
- Services and softwares, such as Matterport, that produce virtual tours using 360-degree still images.

360 still images for 3D visualization

- Geollery (2019)

GSV images for textures of urban models. Integrated with displaying geo-tagged SNS posts.

https://geollery.com/ (CHI2019)

- Avatar360 (2024)

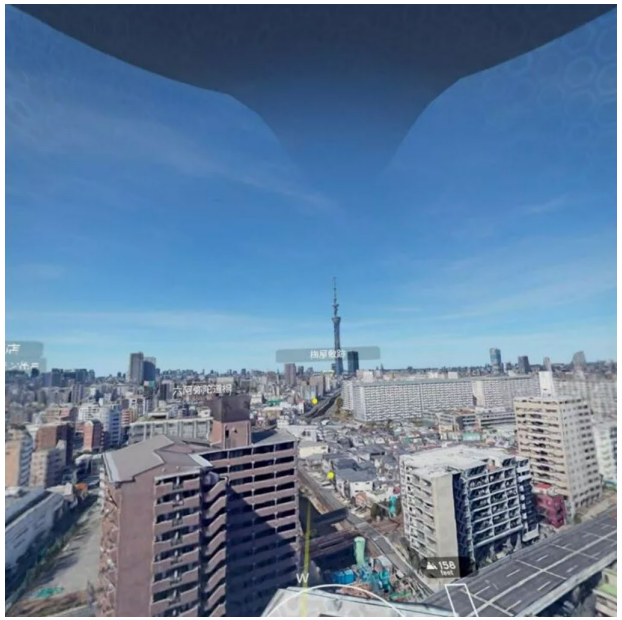
Introducing an avatar for spatially arranged 360° images.

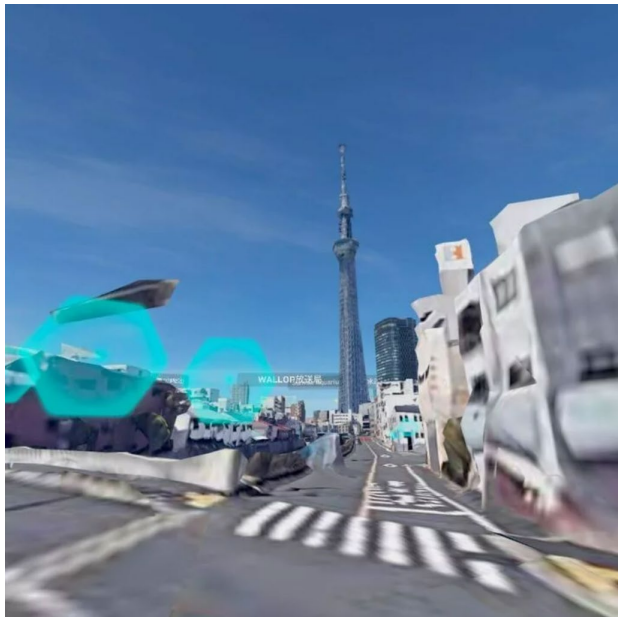
IEEE VR2024

Combination of Google Earth (3D mesh model) and GSV

- FLY (2024) a VR game

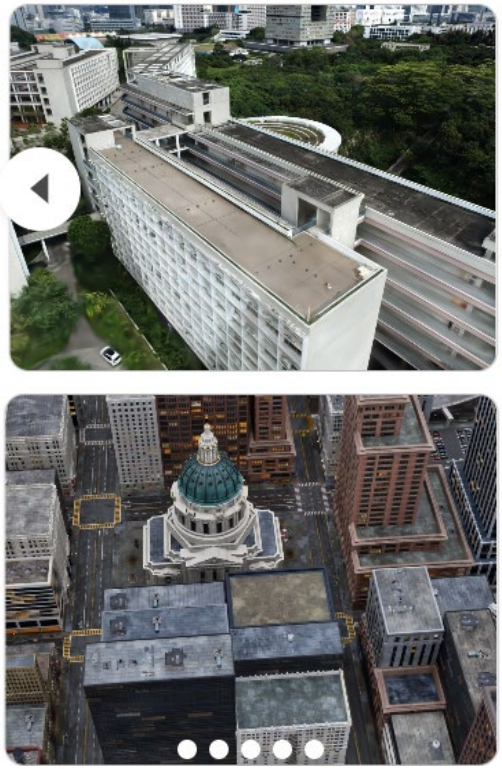
  Fly through with Google Earth and view GSV at ground level.





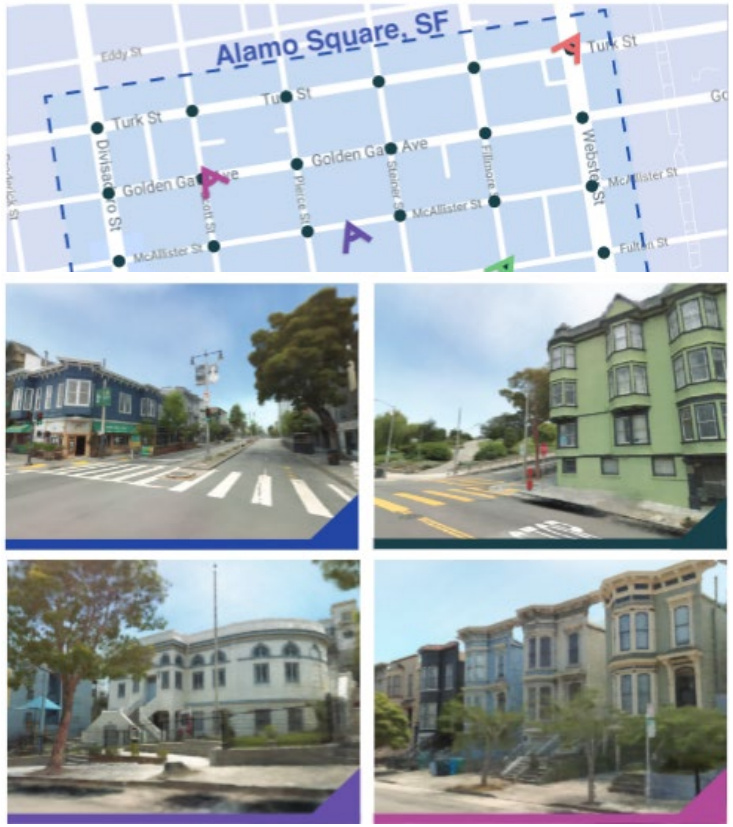
Urban Scale Novel View Synthesis by NeRF or Gaussian Splatting(GS)

- Fly through
  - CityGaussian (ECCV24)



https://dekuliutesla.github.io/citygs/

- Drive/Walk through
  - Block-NeRF (CVPR22)



https://waymo.com/research/block-nerf/

- Entity-NeRF (CVPR24)

https://otonari726.github.io/entitynerf/

Photo-realistic Representation of Urban Areas

Static Image Repr.

- GSV (2007～)
- 3D + GSV texture
  - Geollery (2019)
  - Avatar360 (2024)
  - FLY (2024)

No Motion

Video Repr.

- AspenMovieMap (1980)
- Movie Map (2020)
- 360RVW (2023)
- 360RVW+CityGML (2024)

Motion

NeRF, Gaussian Splatting

Novel view generation from multiview images

Drive/Walk through

- Block-NeRF (2022)
- SUDS (2023)
- Entity-NeRF(2024)

Fly through

- CityNeRF (2022)
- Mega-NeRF (2022)
- CityGaussian(2024)

No Motion

video segment data made from 360° videos in urban areas

Movie Map

- Interactive video that moves around the urban areas
- Accessible from a web browser
- In operation by moviemap.jp

360 RVW

- Pseudo 3D
- Introducing avatars and objects
- Accessible via browser using WebGL
- Extending to smartphone app

360xCityGML RVW

- Use of global city models (CityGML, Plateau)
- Dynamic video mapping
- 3D visualization such as flood etc.

Building Movie Map

N. Sugimoto, Y. Ebine, K. Aizawa, Building Movie Map - A Tool for Exploring in a City - and its Evaluations ACM Multimedia 2020

Movie Map

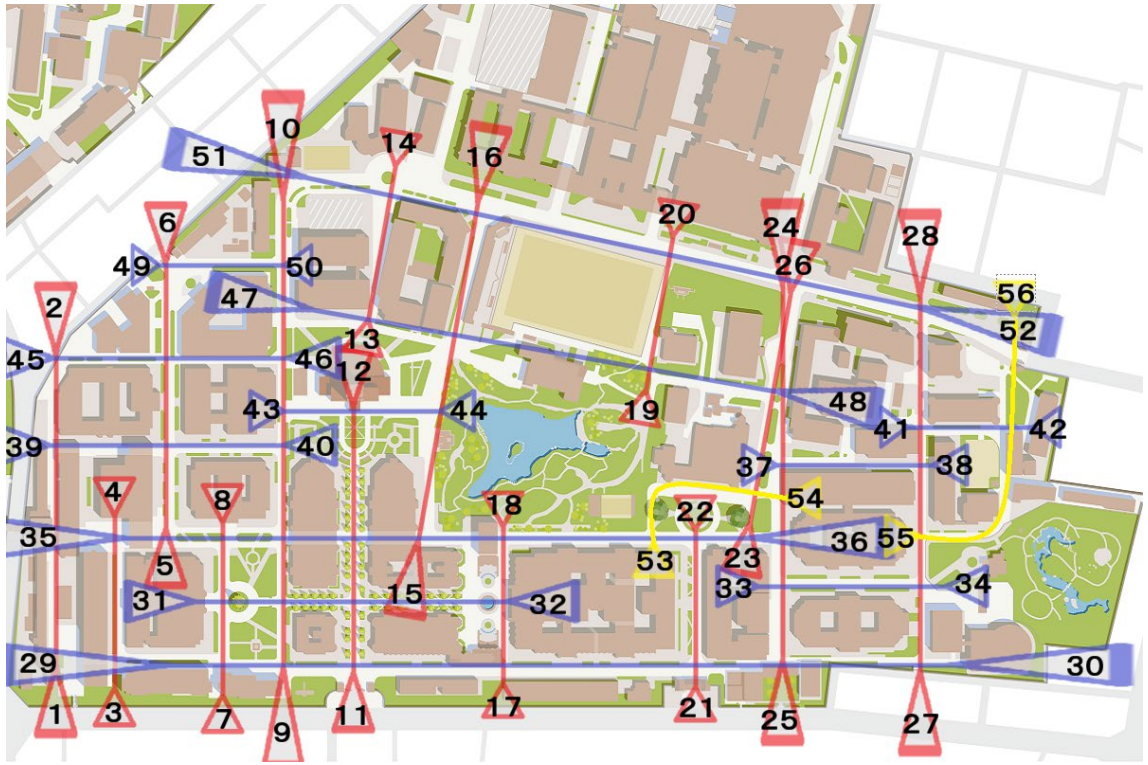
https://moviemap.jp/

demo

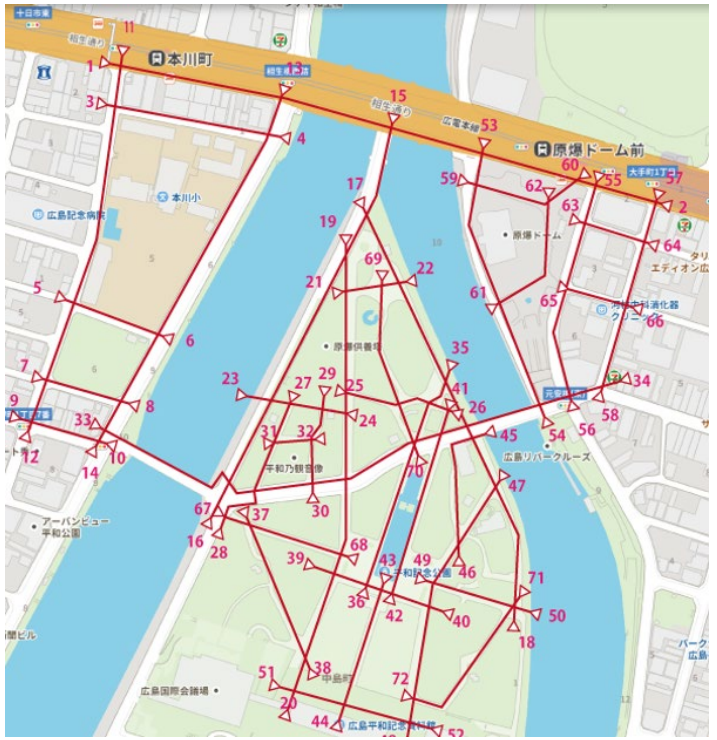
Explore regions by connecting 360° video segments

360° Video Capture

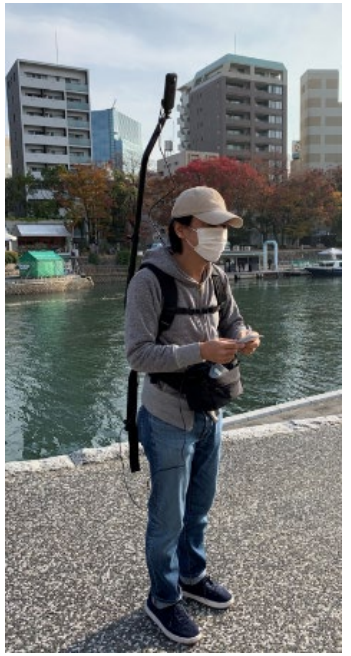
Capture 360° videos while moving along the streets in the area



Univ. of Tokyo Hongo Campus

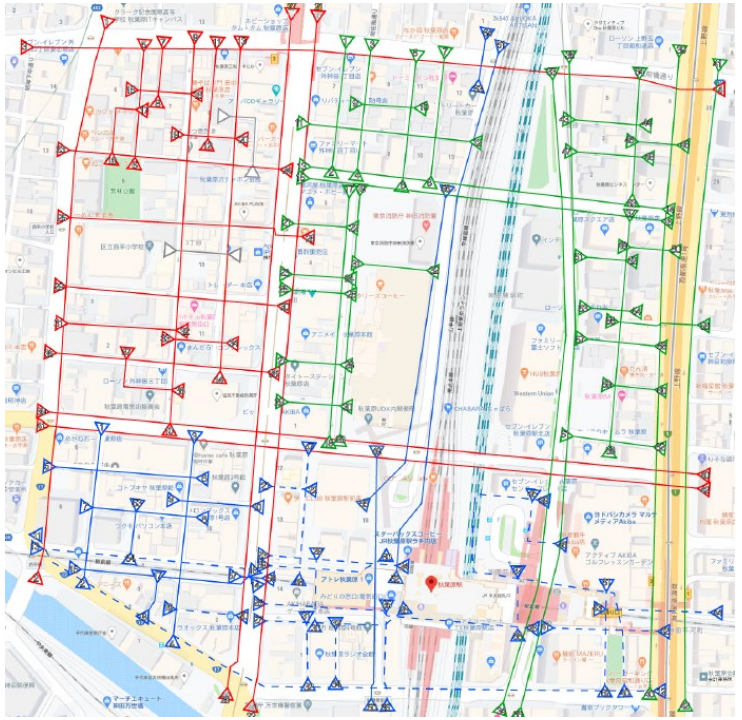


Hiroshima Peace Memorial Park area



Akihabara

https://moviemap.jp



Akihabara 170 videos / 85 streets

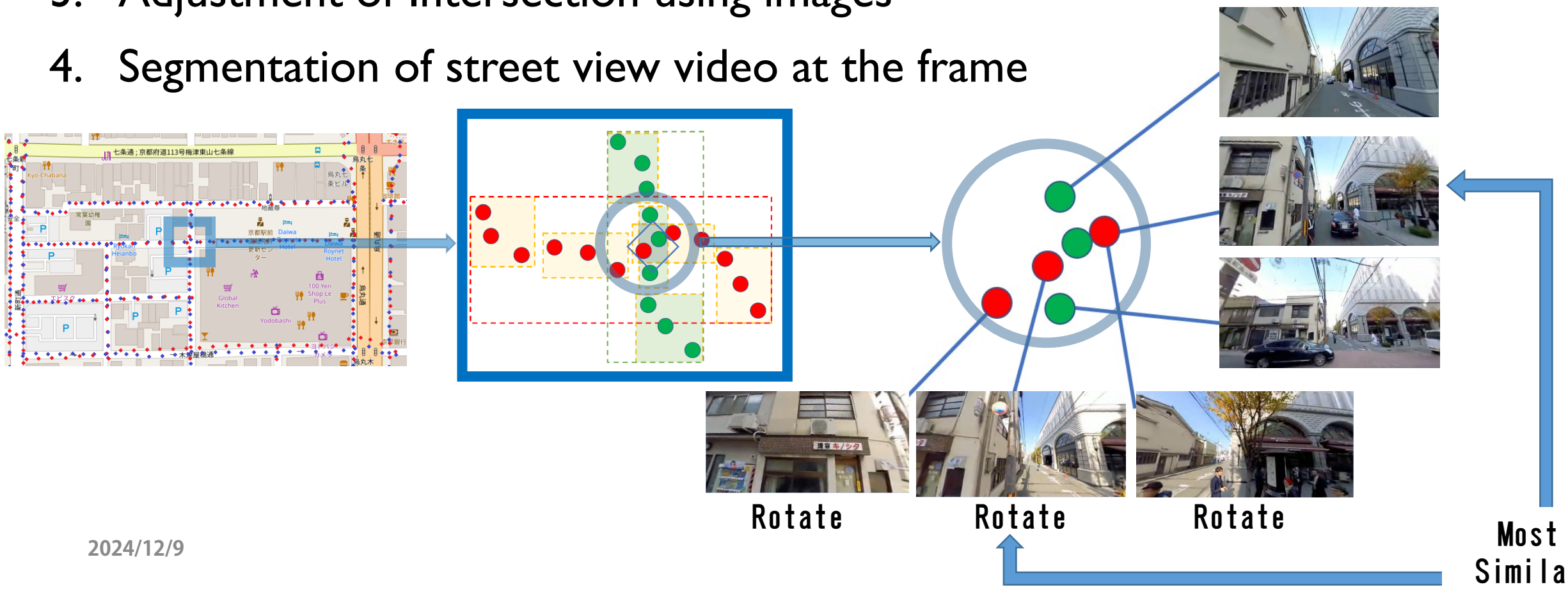
Shimokitazawa 188 videos / 96 streets

Hongo campus 54 videos / 27 streets

YC

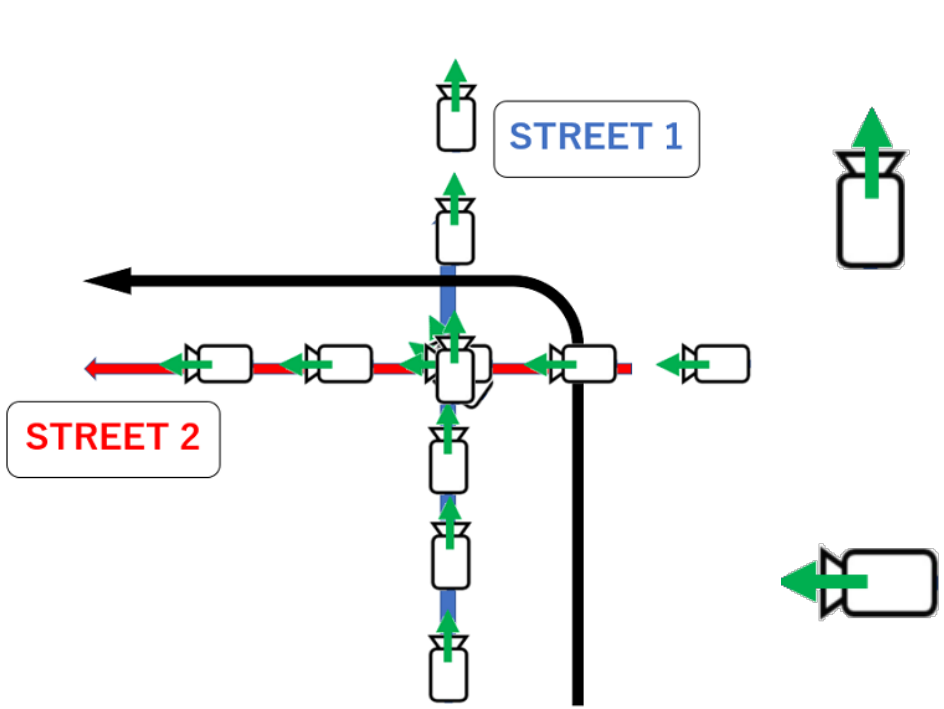
Video Segmentation

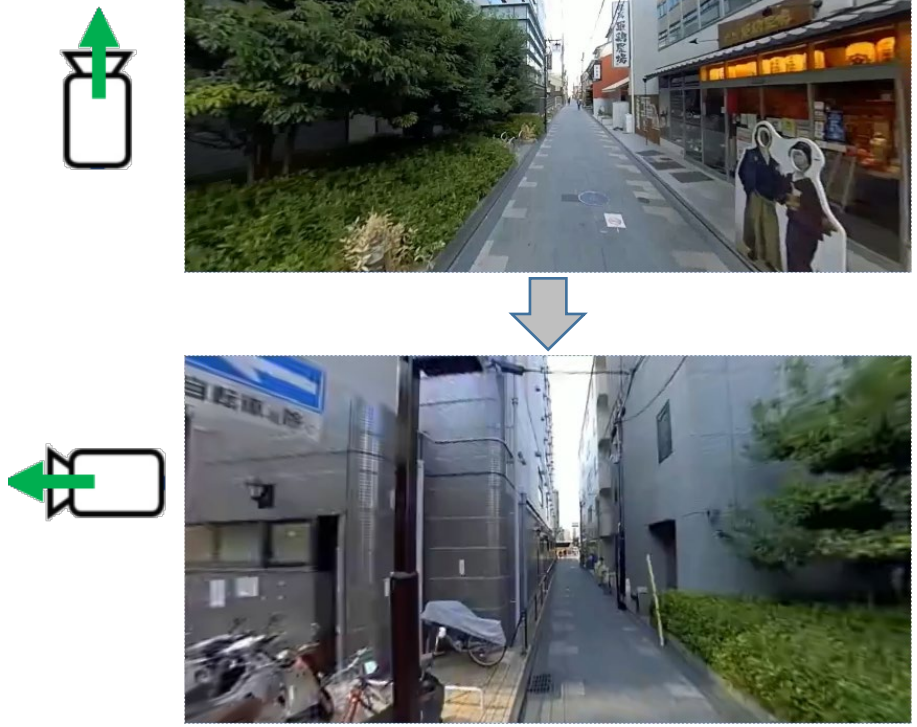
1. Camera location estimation – vSLAM and start/end position
2. Detection of Intersection (a frame pair) using the locations
3. Adjustment of Intersection using images
4. Segmentation of street view video at the frame

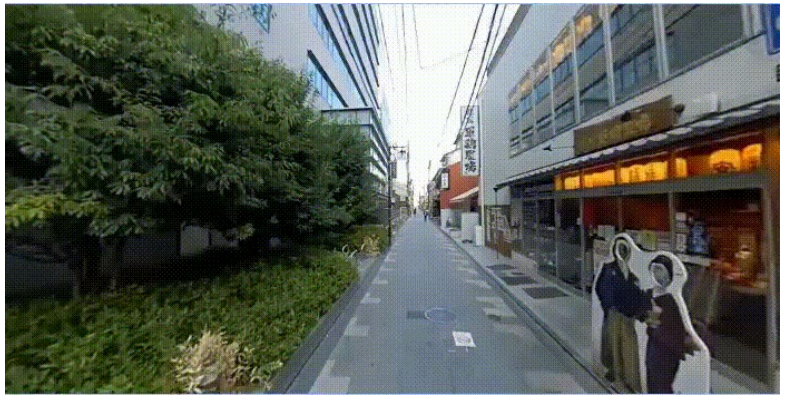


Synthesis of Smooth Turning Views

- Synthesize a turning view by blending the frame pair of the intersection, that makes the video-switching smoother.
- Blending the frames based on the rotation angle.

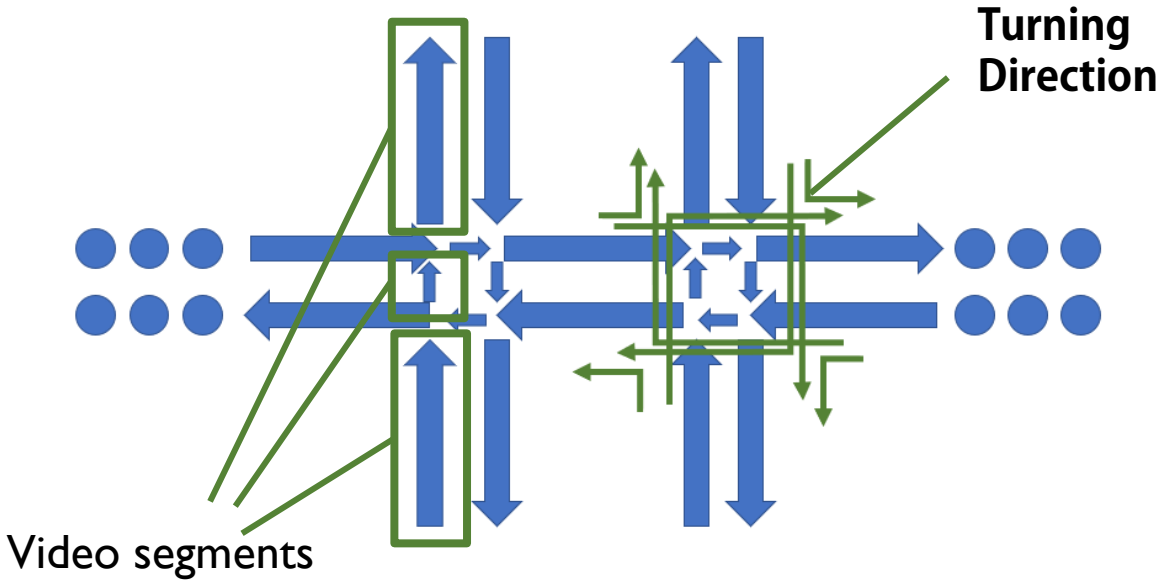






Video Segments and Connections

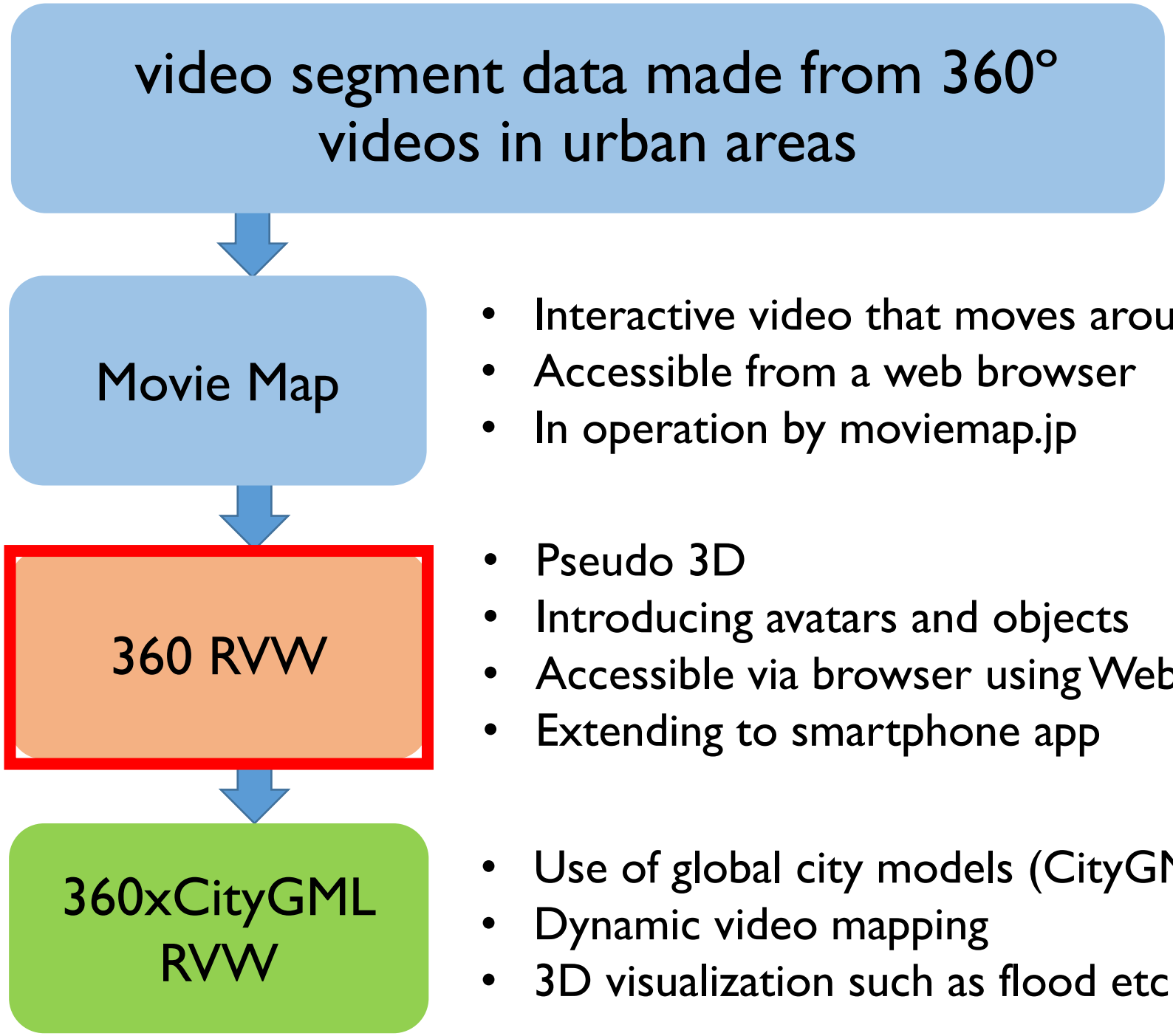
Videos are segmented at the intersections. Video segment can be switched to another at the intersections. Camera locations of the video frames are needed.



Visualization by CAVE Display

- Almost life-size
- Immersive

- With CAVE of Solidray Co. Ltd.



- Interactive video that moves around the urban areas
- Accessible from a web browser
- In operation by moviemap.jp

- Pseudo 3D
- Introducing avatars and objects
- Accessible via browser using WebGL
- Extending to smartphone app

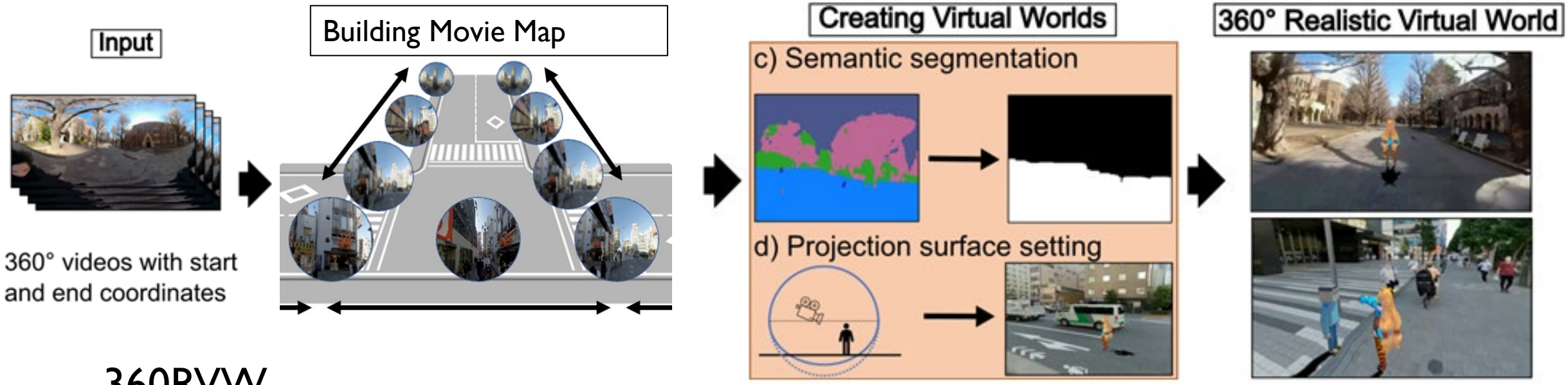
- Use of global city models (CityGML, Plateau)
- Dynamic video mapping
- 3D visualization such as flood etc.

360RVW

Fusing 360° Real Video and Interactive Virtual World

M. Takenawa, N. Sugimoto, L. Wöhler, S. Ikehata, K. Aizawa
360RVW: Fusing Real 360° Videos and Interactive Virtual Worlds,
ACM Multimedia 2023

360RVW



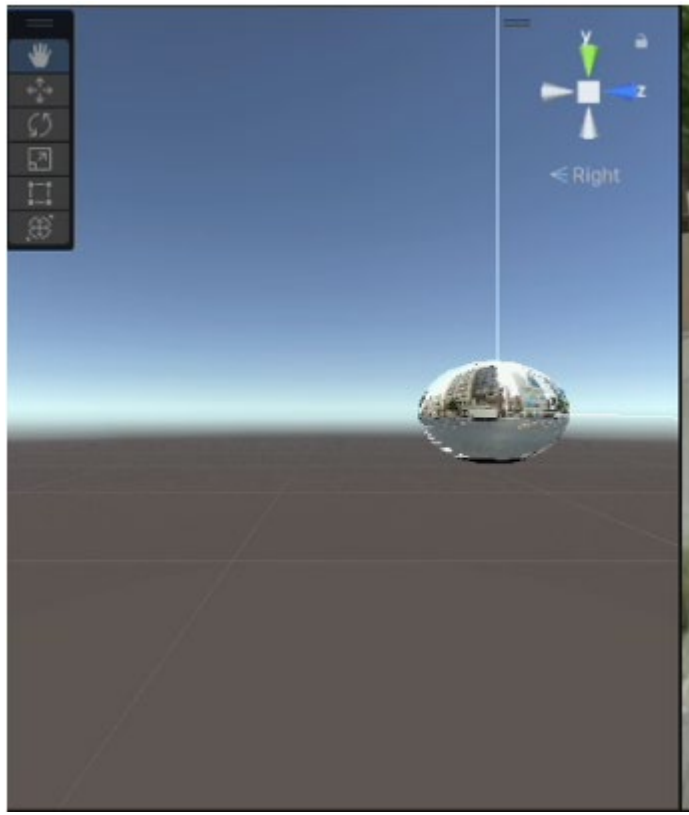
360RVW

- Video streaming from the Movie Map database.
- Virtual world made by projection of video onto a sphere-like surface.
- Rendering avatars and objects inside the surface.
- Use of a semantic map constraining avatar's movement.

Inside and outside of the projection surface

- The virtual world is a projection onto the surface.
- The projection changes according to the location of the viewer (avatar). The viewer is always inside.

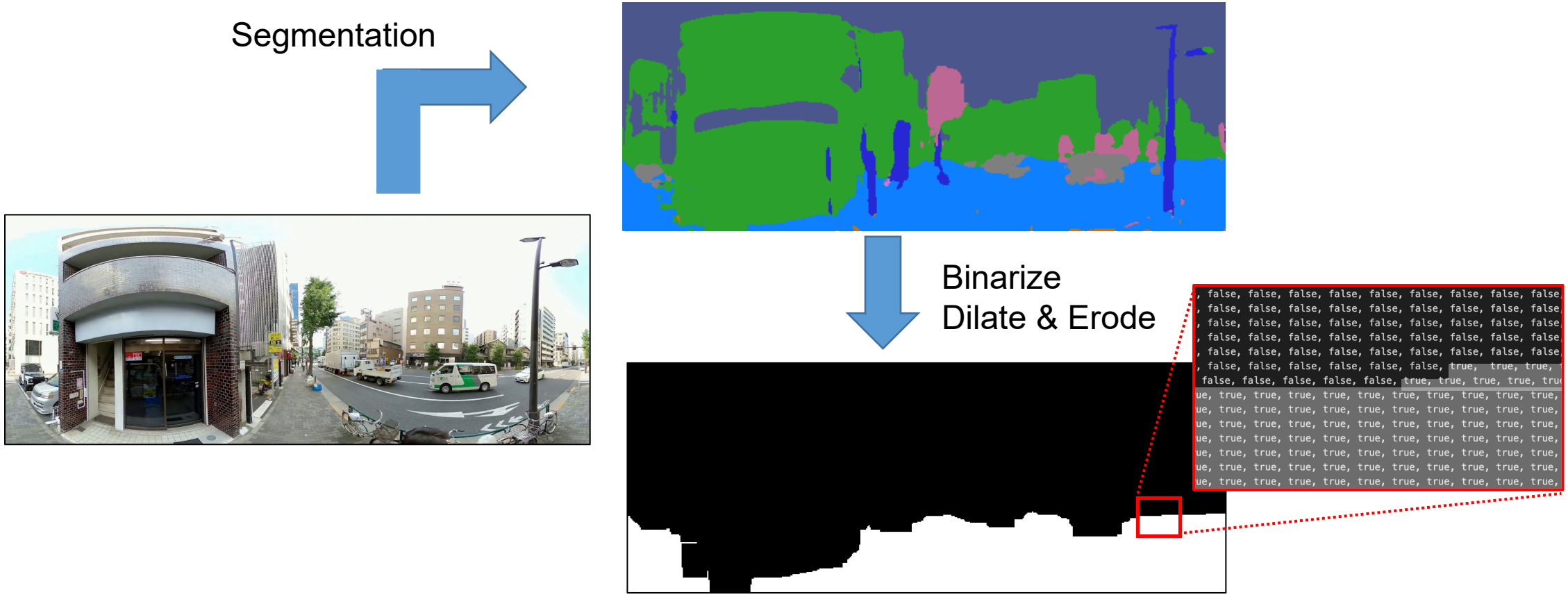
OUT



IN

Semantic Map

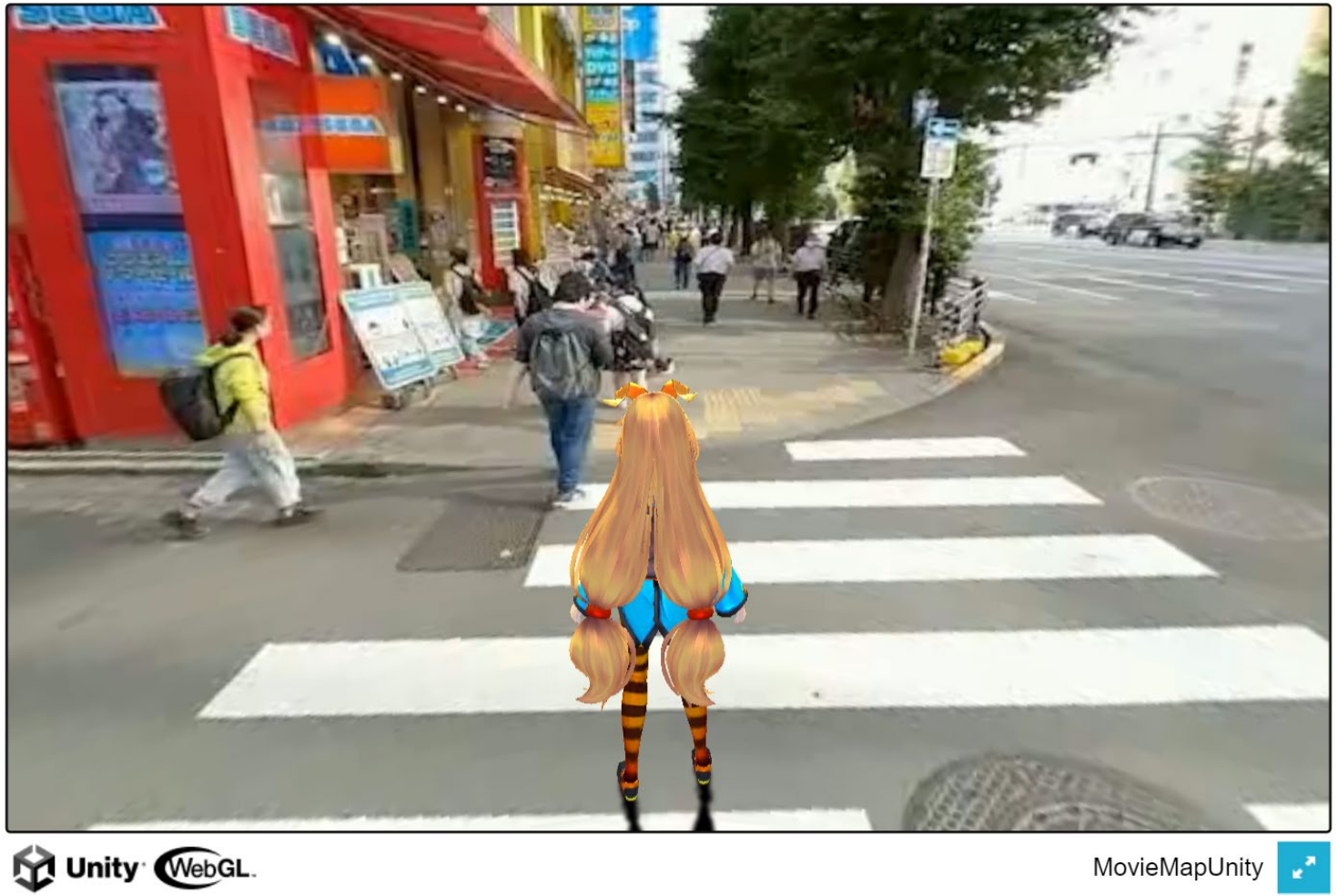
- Semantic segmentation applied to frames to determine “road”.
- The avatar’s moving area is constrained by it.



Demo : 360RVW for PC

Demo

WebGL: Accessible from Web



Demo WGL

https://moviemap.jp/unity/MovieMapUnity/

Smartphone App

Demo

iPhone

video segment data made from 360° videos in urban areas

Movie Map

- Interactive video that moves around the urban areas
- Accessible from a web browser
- In operation by moviemap.jp

360 RVW

- Pseudo 3D
- Introducing avatars and objects
- Accessible via browser using WebGL
- Extending to smartphone app

360xCityGML RVW

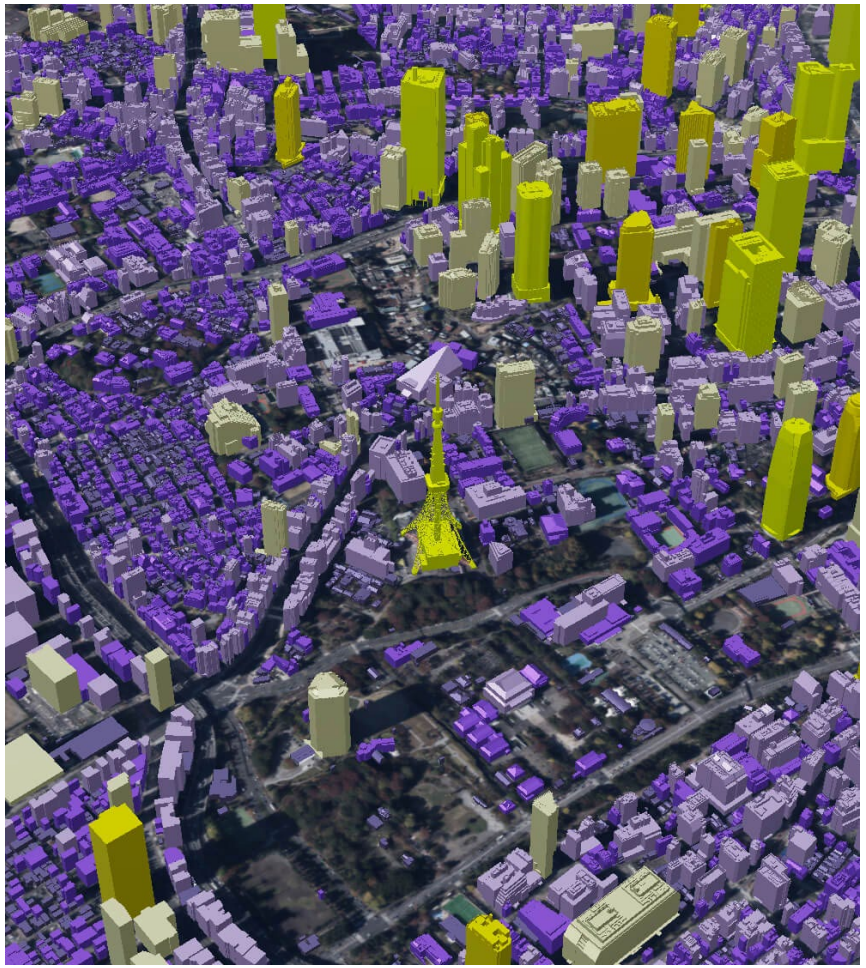
- Use of global city models (CityGML, Plateau)
- Dynamic video mapping
- 3D visualization such as flood etc.

360° Video x City Model (Plateau) for Building Virtual World

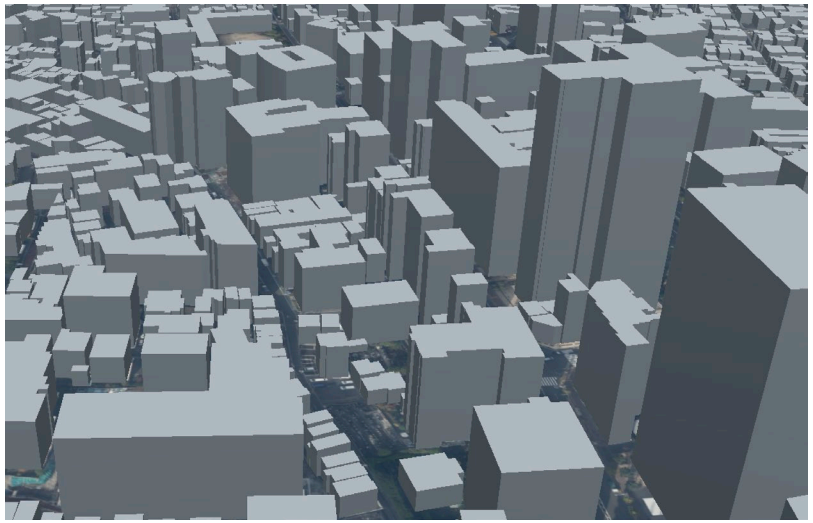
Plateau Innovation Award 2024.2

Plateau

- Plateau is a project led by the Ministry of LIT that develops 3D urban models across Japan and release them as open data.
- Not only 3D city models but also various geo-spatial information are included.
- CityGML format data and various conversions are available.



Plateau City Model (LOD 1)



https://plateauview.mlit.go.jp/

360RVW	PLATEAU
❌ 3D City Model	✅ 3D City Model
✅ Realism	❌ Realism

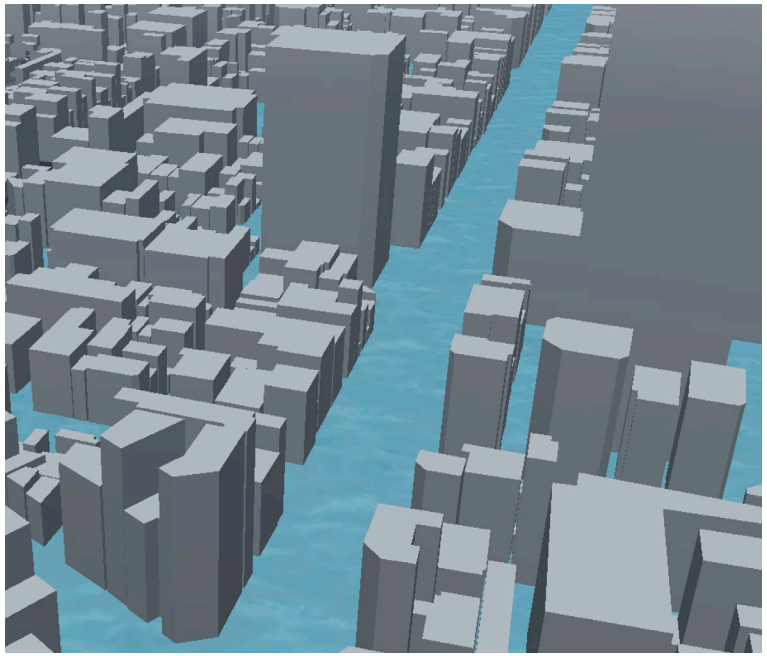
360° Video x Plateau for building virtual world.

Virtual World by 360° Video x Plateau

https://www.openstreetmap.org/copyright

Projection of 360° Video onto the City Model

Flood Visualization



https://plateauview.mlit.go.jp/

PLATEAU's 3D flooding risk data

https://www.openstreetmap.org/copyright

Flood visualization in the virtual world

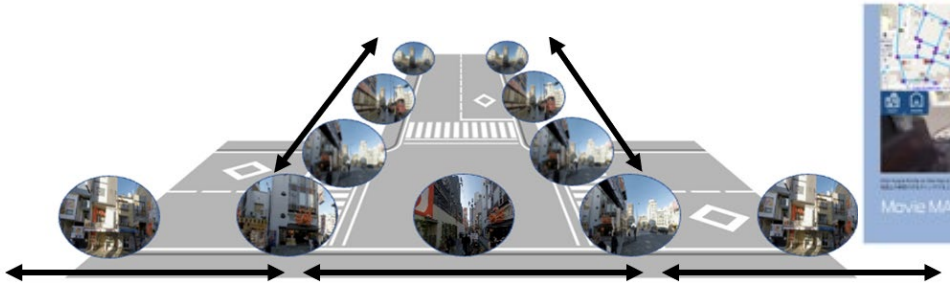
Connecting Geospatial Information with Real World Contexts

Real 360° video is efficient for building realistic virtual world.

360° Video Database

- Movie Map
- 360 RVW
- 360xPlateau RVW

- Interactive video



- Pseudo 3D with avatars

360° video segments

- Global 3D

360° videos, city models

Thank you for your attention.

Our team:
N. Sugimoto, M. Takenawa, T. Banno, L. Woehler, S. Ikehata,
T. Yoshida, K. Suzuki