Mittuniversitetet
MID SWEDEN UNIVERSITY

Challenge-Driven Quality Evaluation and Dataset Development for Immersive Visual Experiences

Emin Zerman, PhD – Associate Senior Lecturer (~Assistant Professor)

Mid Sweden University

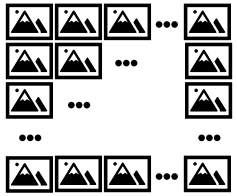
Spring School on Social XR 2025 | Amsterdam | 2025-04-09

(Immersive) Telepresence

Internet

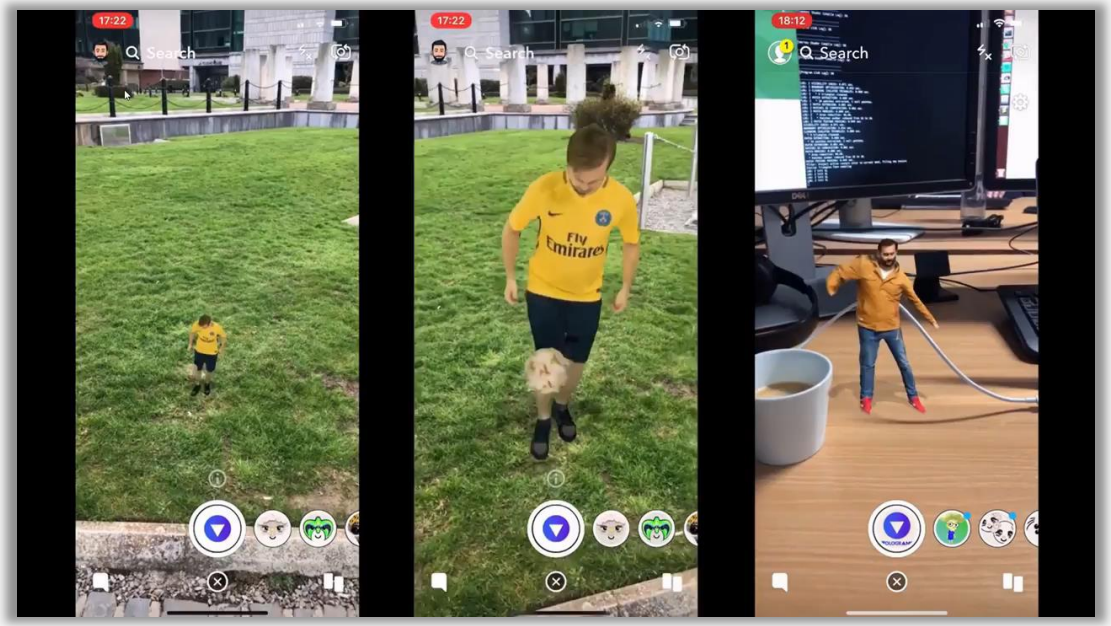
Volumetric video

Internet

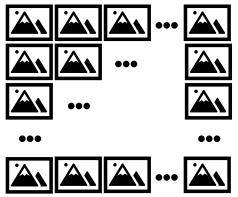


Light fields

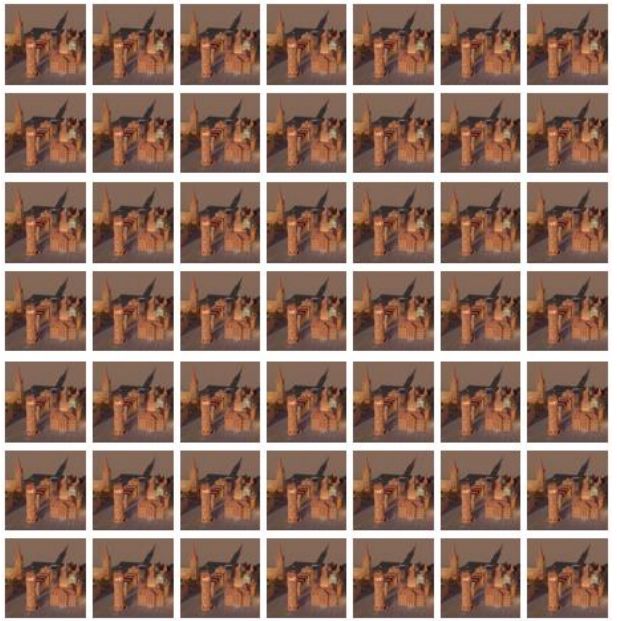
Volumetric video



Videos courtesy of Aljosa Smolic & Volograms



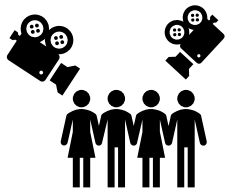
Light fields



Front-to-back

Back-to-front

BASICS PCQA dataset



Volumetric video

VSENSE-VVDB2

Motivation

Limited number of content

No learning-based compression

Contained only/mostly humans

Challenge #1

Few source contents

Few semantic categories

Building BASICS – Collecting interesting models

BASICS Database

Requirements

- Interesting content for telepresence & telecommunication applications
- Available to us – *or publicly available*
- Can be made publicly available afterwards – *redistribution*

Challenge #1

Few source contents

Few semantic categories

VSENSE | **XD Productions** | **???**

Ak, A., Zerman, E., Quach, M., Chetouani, A., Smolic, A., Valenzise, G., & Le Callet, P. (2024). **BASICS: Broad quality assessment of static point clouds in a compression scenario.** *IEEE Transactions on Multimedia*.

Building BASICS – Collecting interesting models

BASICS Database

VSENSE

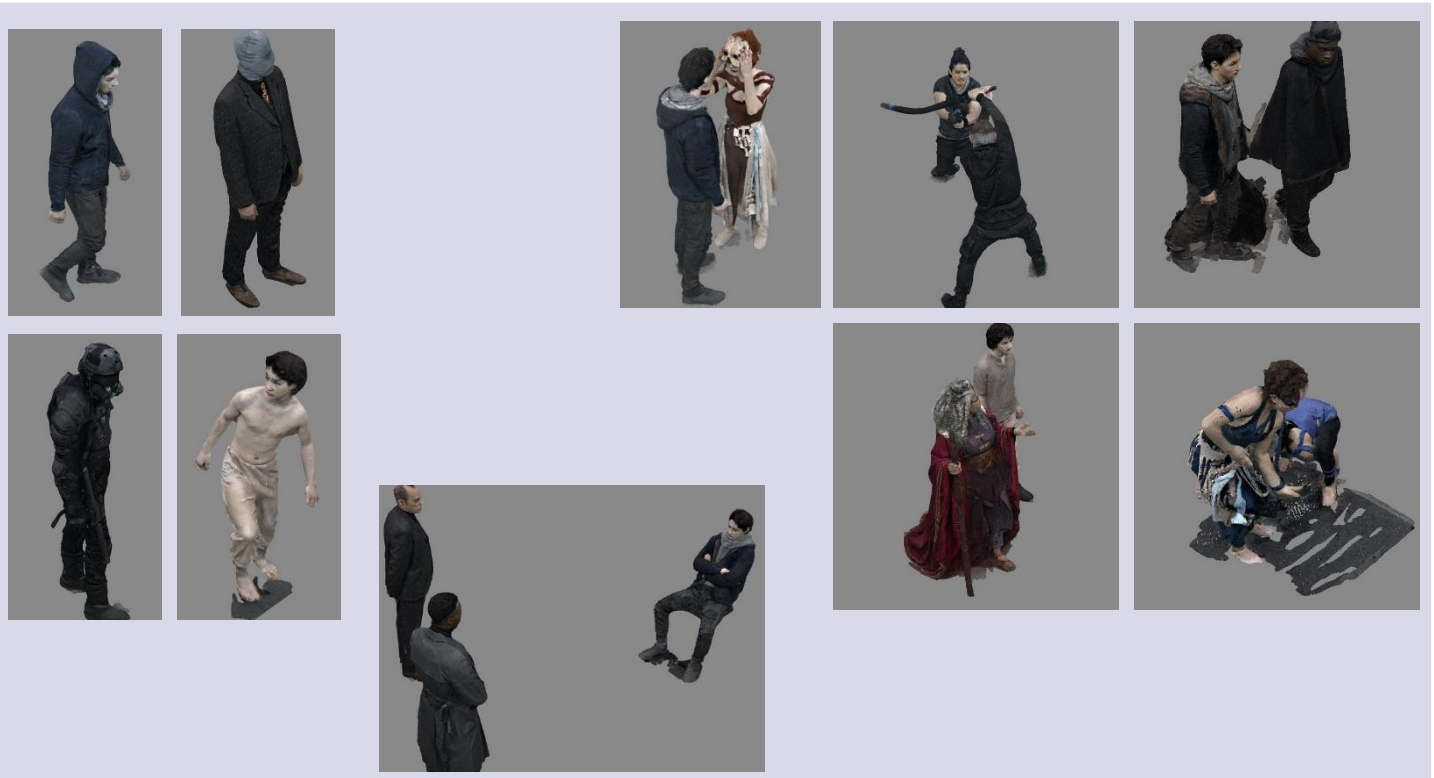
XD Productions

???

Building BASICS – Collecting interesting models

BASICS Database

XD Productions



VSENSE

???

Building BASICS – Collecting 3D models

BASICS Database

Requirements

- Interesting content for telepresence & telecommunication applications
- Available to us – *or publicly available*
- Can be made publicly available afterwards – *redistribution*

VSENSE

XD Productions

???

People

Sketchfab

3D models with CC-BY or CC-0 licences

Sketchfab EXPLORE BUY 3D MODELS FOR BUSINESS UPLOAD

Chateau de la Bretesche -RAWscan

3D Model

Andrea Spognetta (Spogna) FOLLOW

Download 3D Model + Add To </> Embed Share

Triangles: **1.3M** Vertices: **632k** More model information

Chateau de la Bretesche

Photogrammetry from video

License: CC Attribution-NonCommercial Learn more

Chateau de la Bretesche -RAWscan
This model is free

Licensing information

CC Attribution-NonCommercial
Author must be credited. No commercial use.
LEARN MORE

Credit the Creator
Chateau de la Bretesche -RAWscan by Andrea Spognetta (Spogna) is licensed under Creative Commons Attribution-NonCommercial
COPY CREDITS

Available downloads

obj Original format		
.obj	49MB	DOWNLOAD
USDZ Converted format		
.usdz	34MB	DOWNLOAD
glTF Converted format		
.gltf	46MB	DOWNLOAD
GLB Converted format		
.glb	61MB Texture size: **8k**	DOWNLOAD
.glb	50MB Texture size: **4k**	DOWNLOAD
.glb	43MB Texture size: **1k**	DOWNLOAD

Use Importers to search for 3D models directly from your favorite software.

Need a more flexible license and high-quality content?
Explore the Sketchfab Store

Mittuniversitetet
MID SWEDEN UNIVERSITY

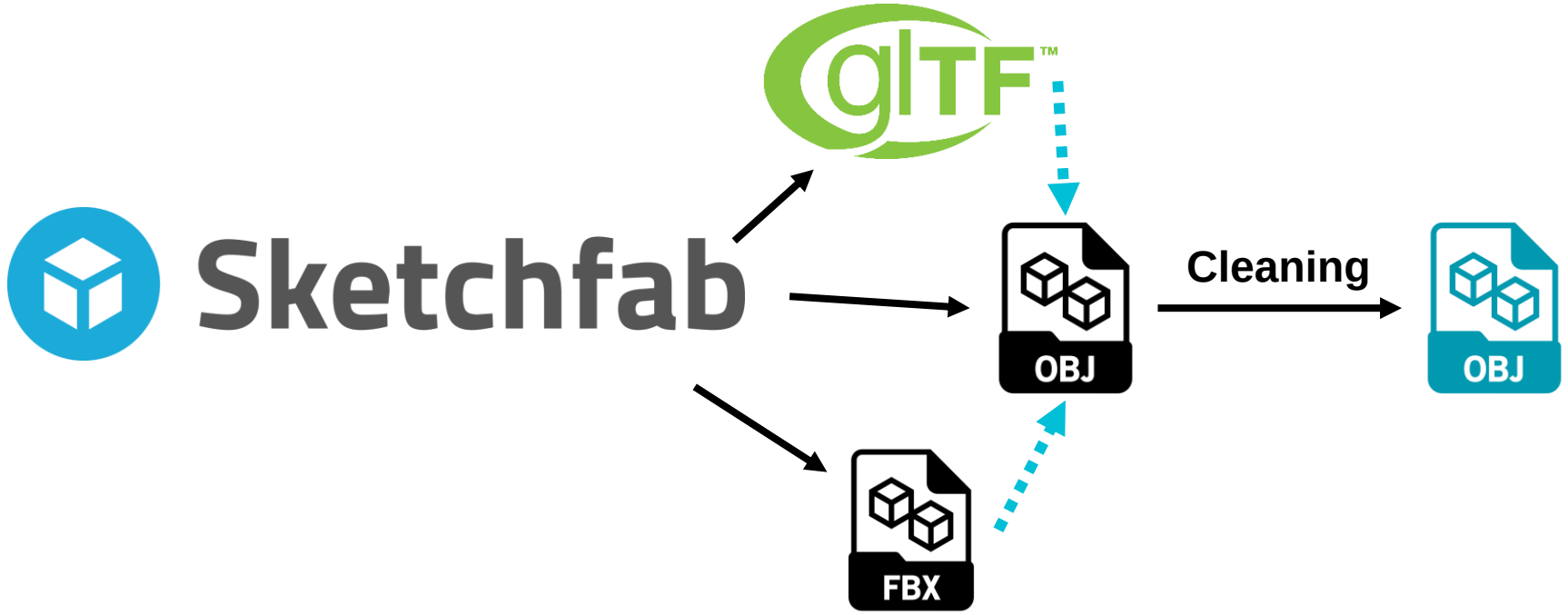
Building BASICS – Collecting 3D models

BASICS Database

VSENSE

XD Productions

Publicly available 3D models w/ CC-BY or CC-0 licence



Building BASICS – Collecting 3D models

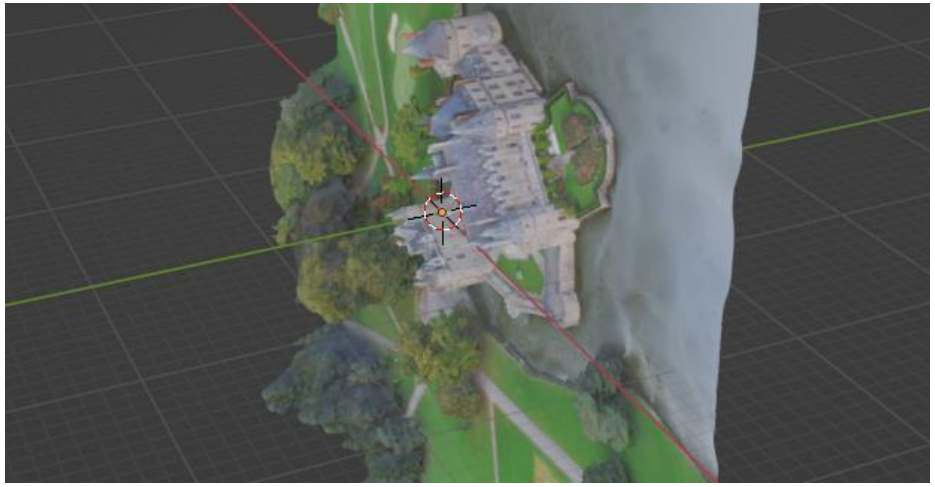
BASICS Database

VSENSE

XD Productions

Publicly available 3D models w/ CC-BY or CC-0 licence

Sketchfab → OBJ → Cleaning → OBJ



Building BASICS – Collecting 3D models

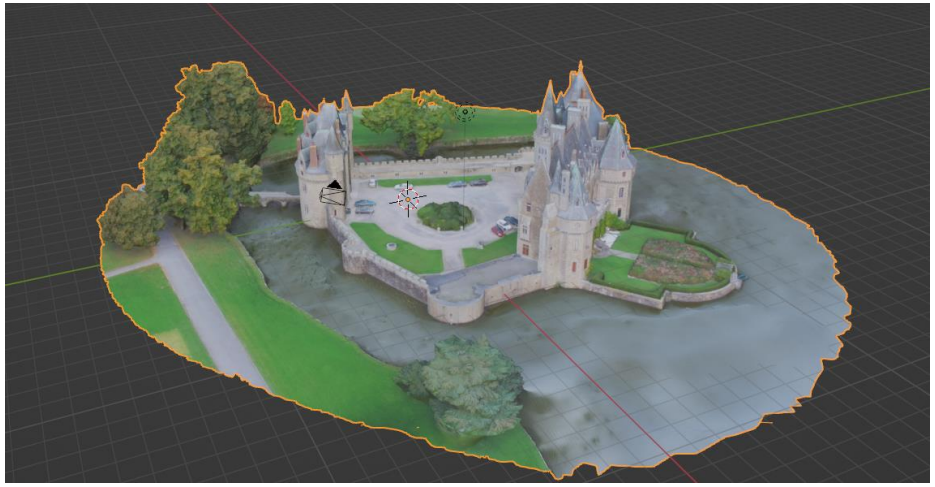
BASICS Database

VSENSE

XD Productions

Publicly available 3D models w/ CC-BY or CC-0 licence

Sketchfab → OBJ → Cleaning → OBJ



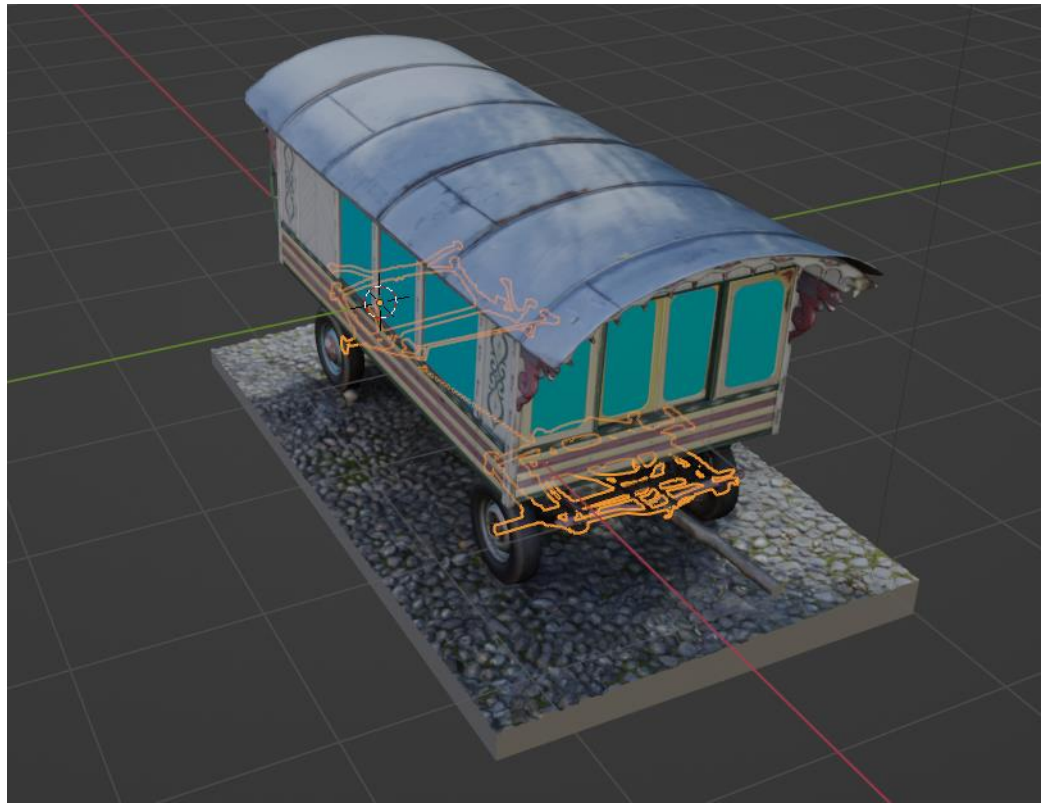
Building BASICS – Collecting 3D models

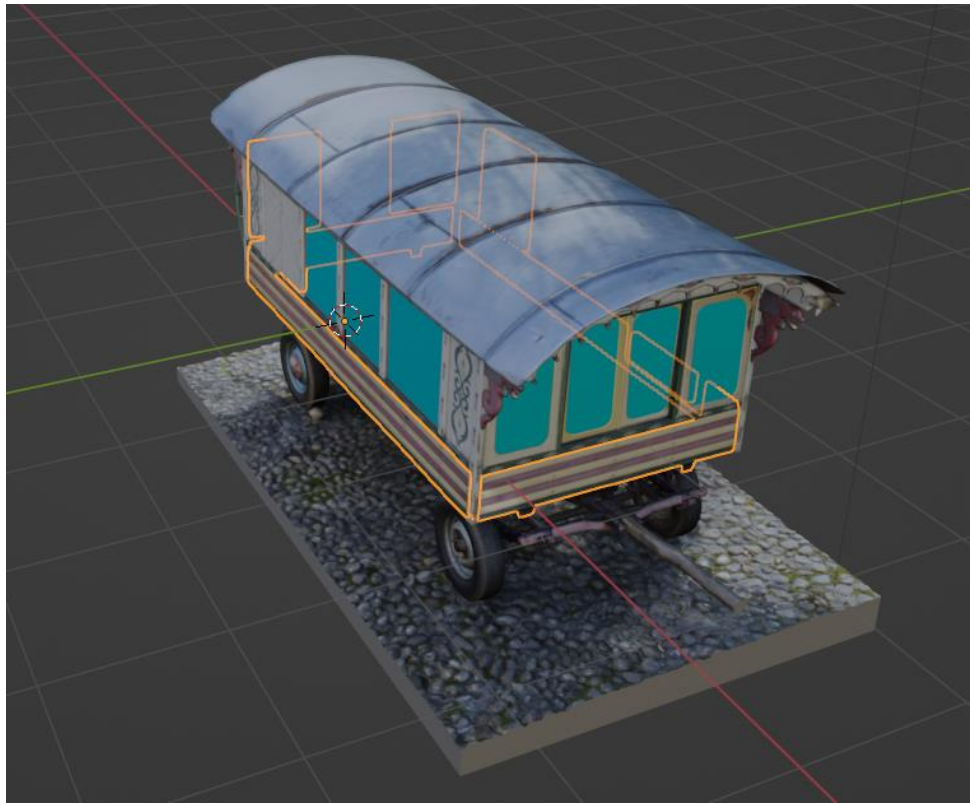
BASICS Database

VSENSE

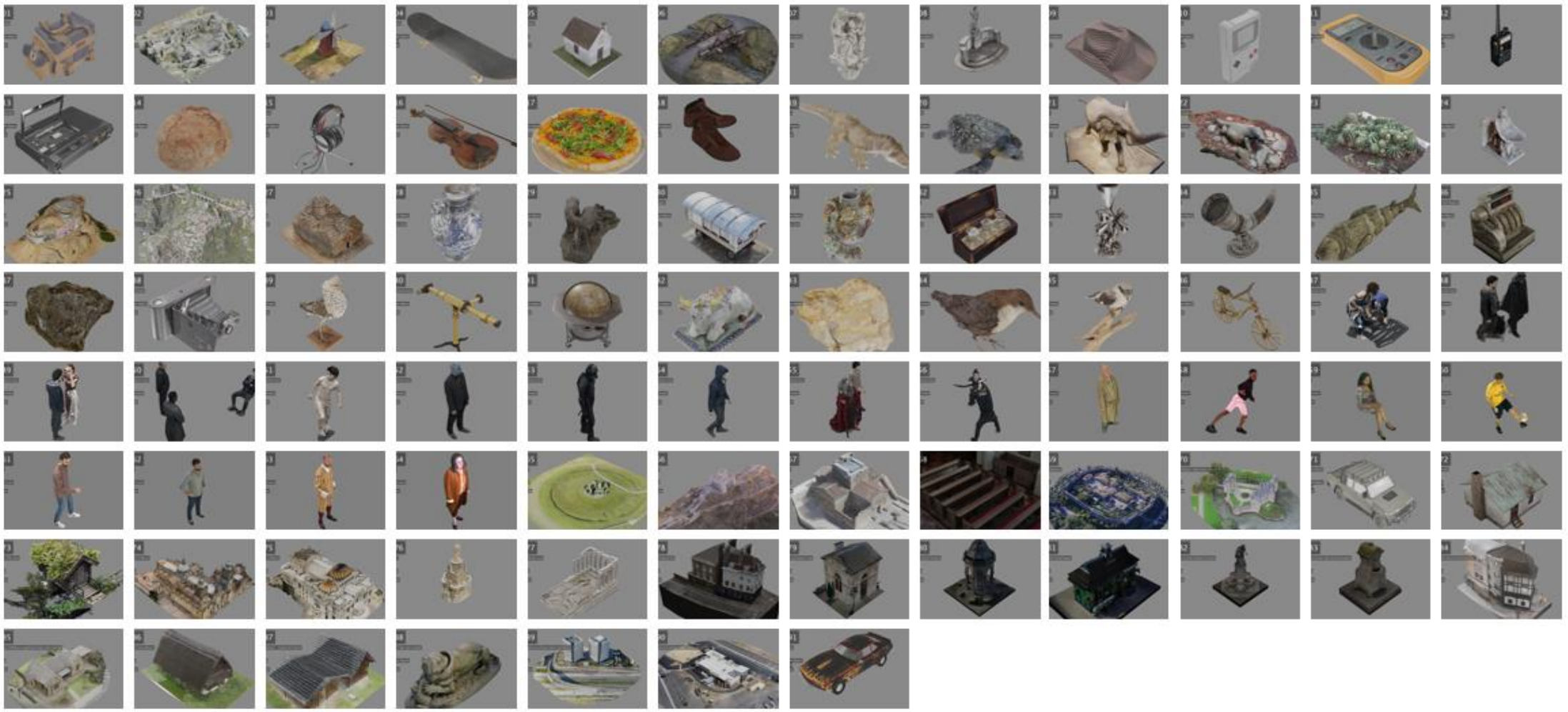
XD Productions

Publicly available 3D models w/ CC-BY or CC-0 licence



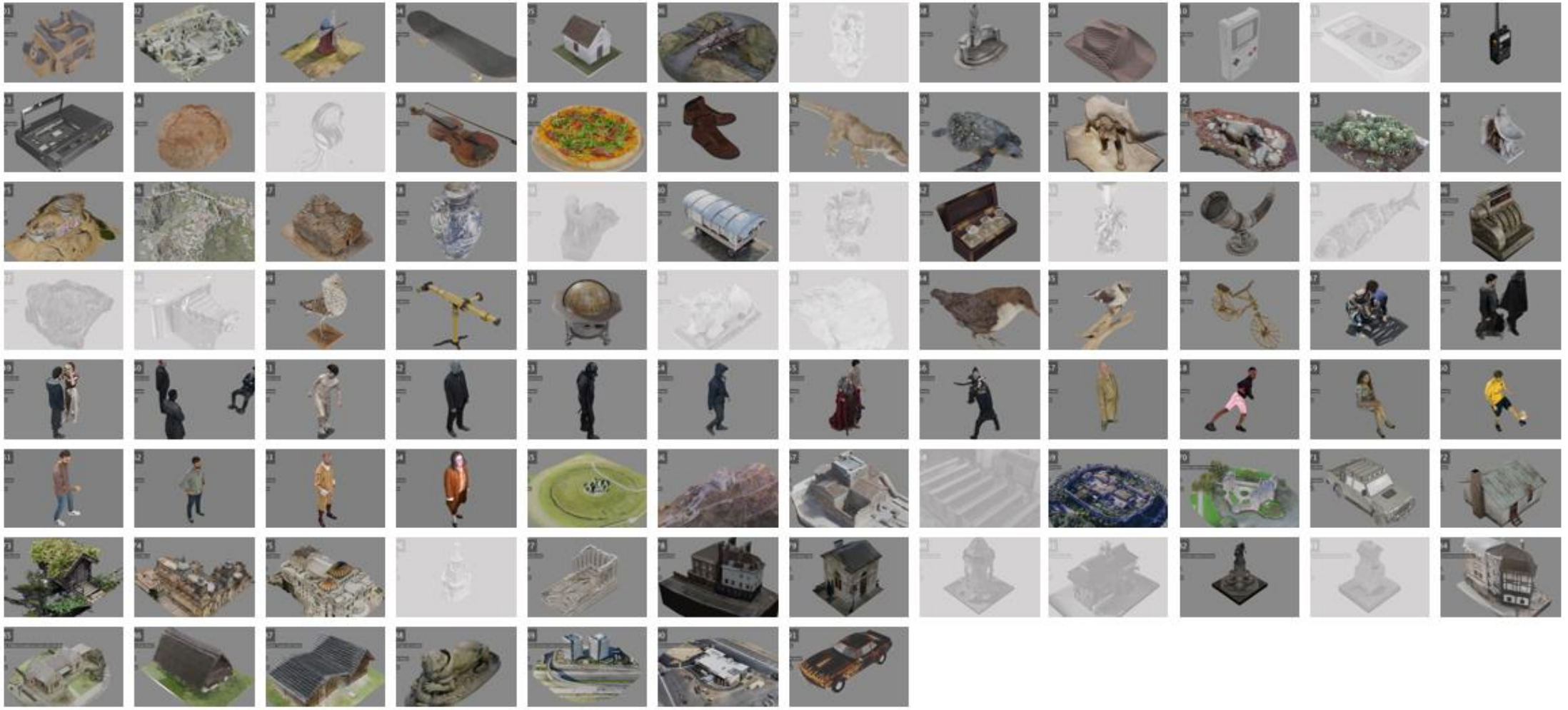


Initial collection – 91 models



BASICS Database

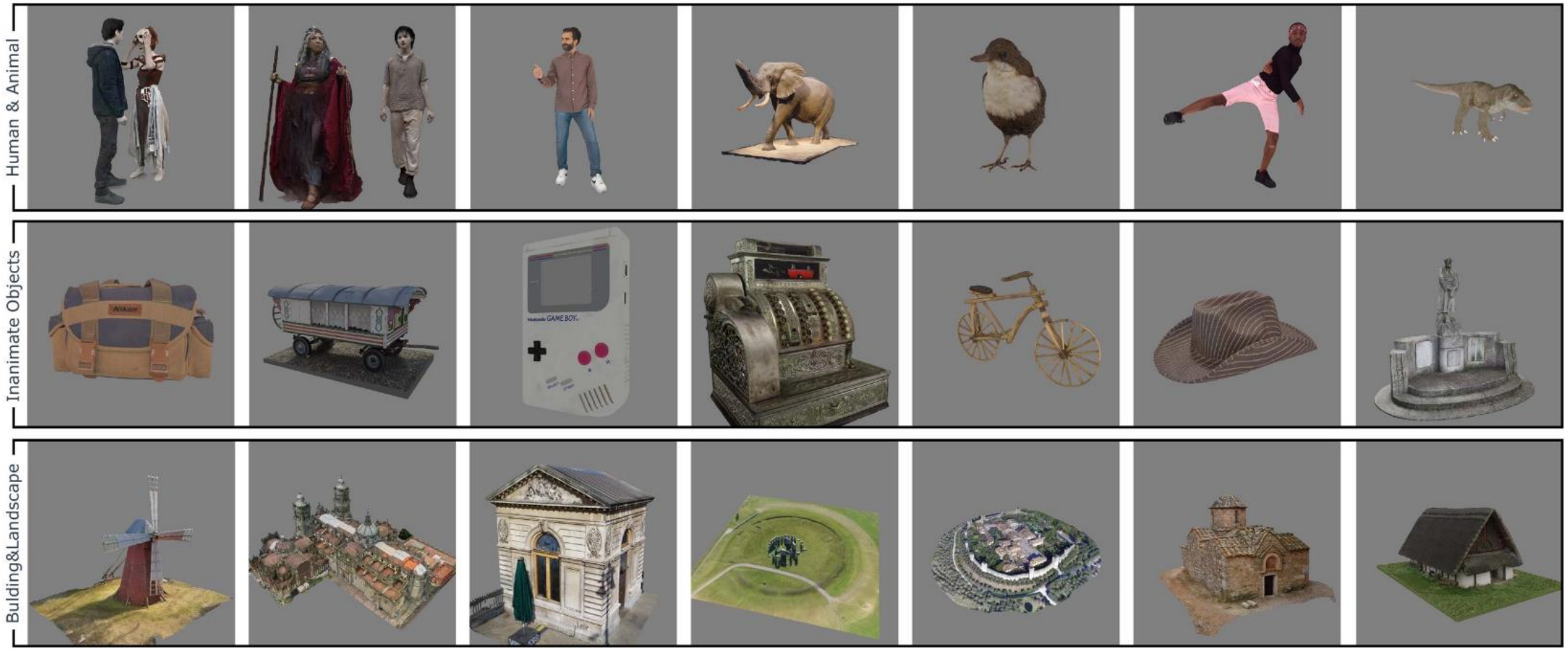
Initial collection – 91 models



BASICS Database

Selected 75 point clouds (visual not exhaustive)

BASICS Database



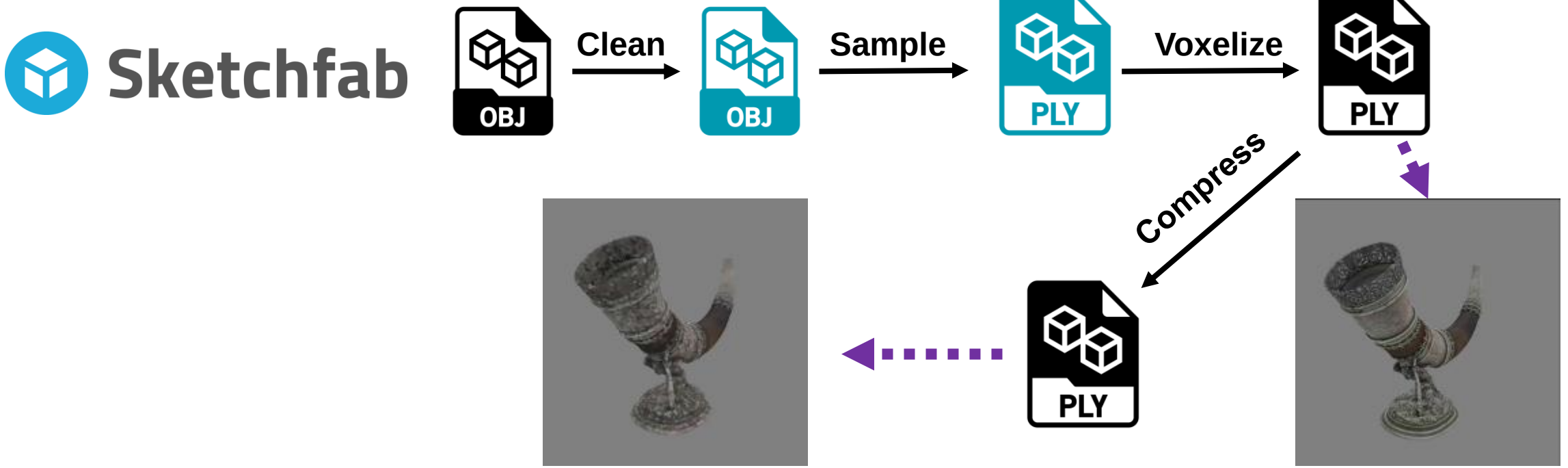
Point cloud processing chain

VSENSE

XD Productions

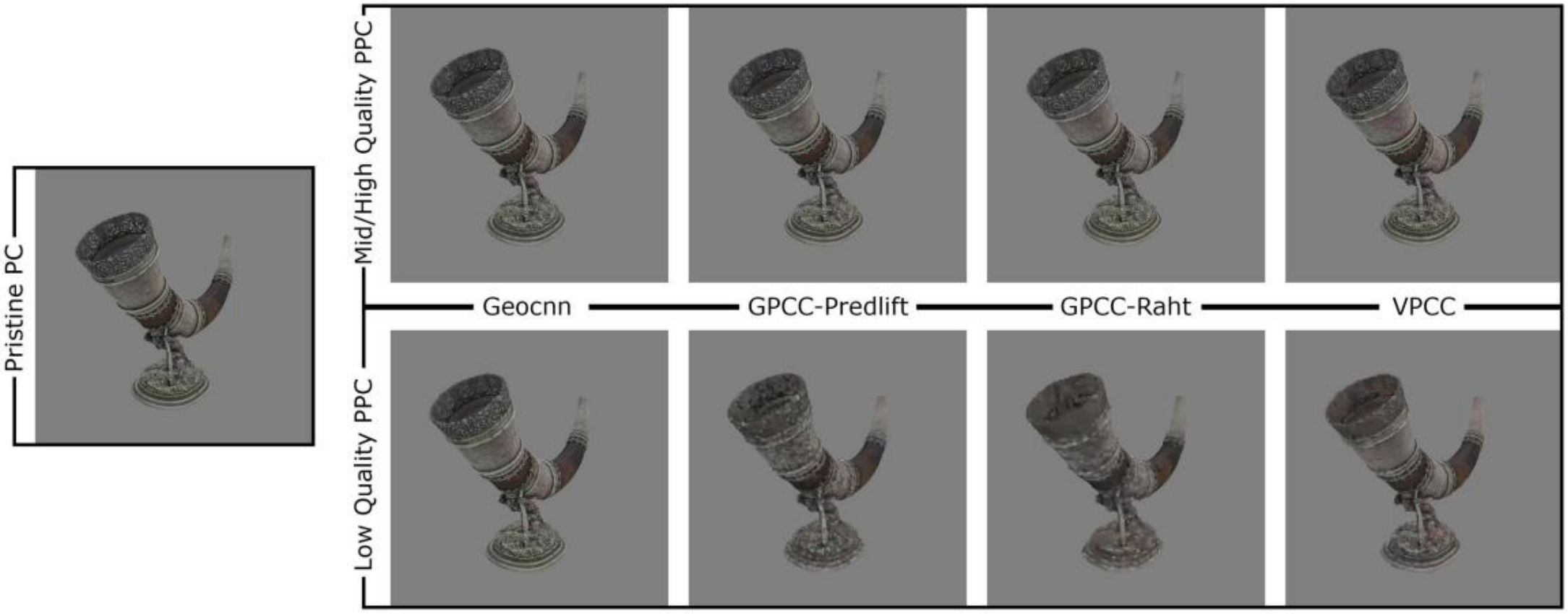
Publicly available 3D models w/ CC-BY or CC-0 licence

BASICS Database



Point cloud processing chain

BASICS Database



BASICS Dataset in ICIP 23 Grand Challenge

BASICS Database

BASICS Database Zenodo

https://doi.org/10.5281/zenodo.8324545

Challenge #2

How to evaluate metrics developed for different tasks

How best to identify “*good*” metrics for our ICIP23 grand challenge

Motivation

Provide a methodology for ecological validity

Provide a toolkit to the scientific community

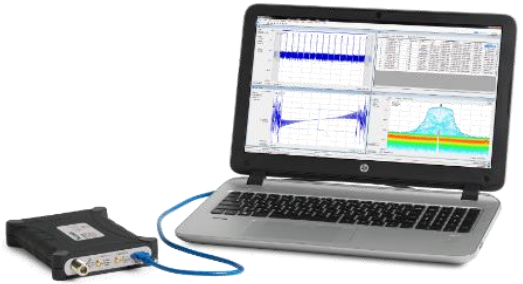
Track-based evaluation

BASICS PCQA dataset

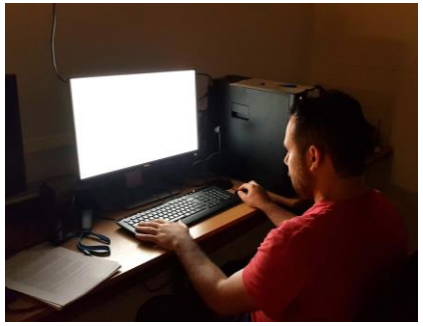
Volumetric video

Ecological Validity

- Ecological validity
  - realism with which a design of evaluation setup matches the user's real work context.
  - making the results applicable to everyday environments and situations.

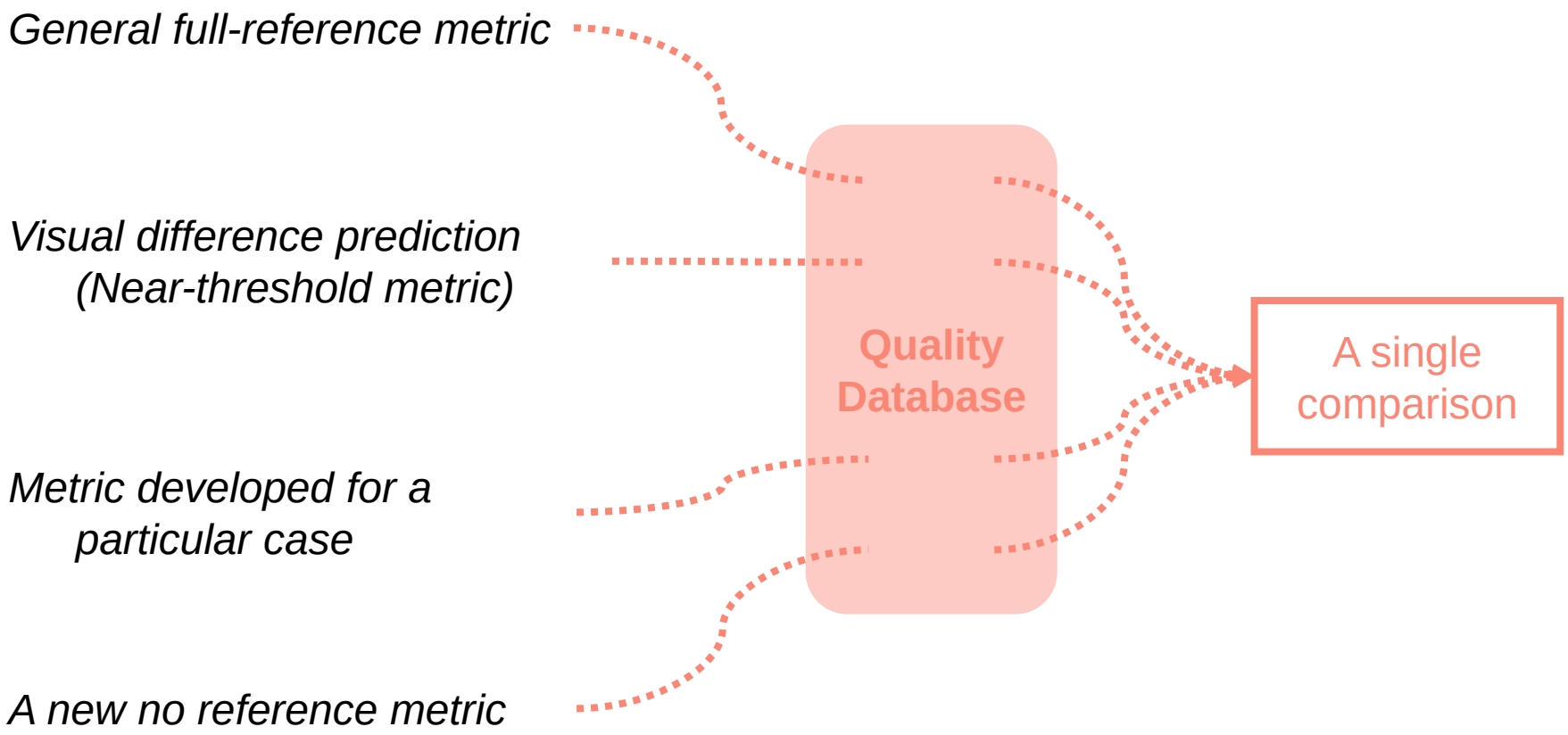


- The closer the study's methods and environment are to real-life conditions, the higher the ecological validity.

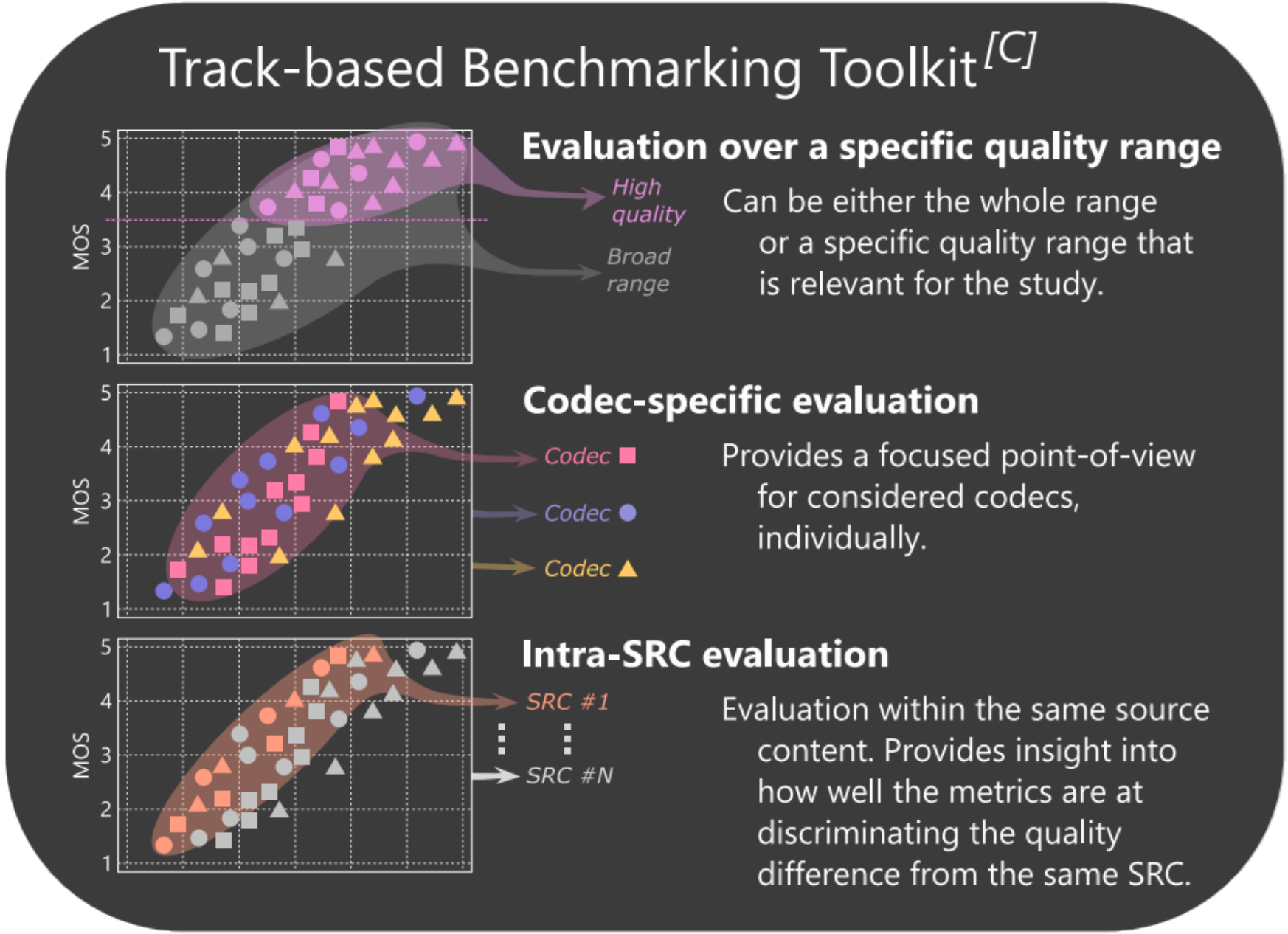


Track-based eval.

Track-based Evaluation Methodology



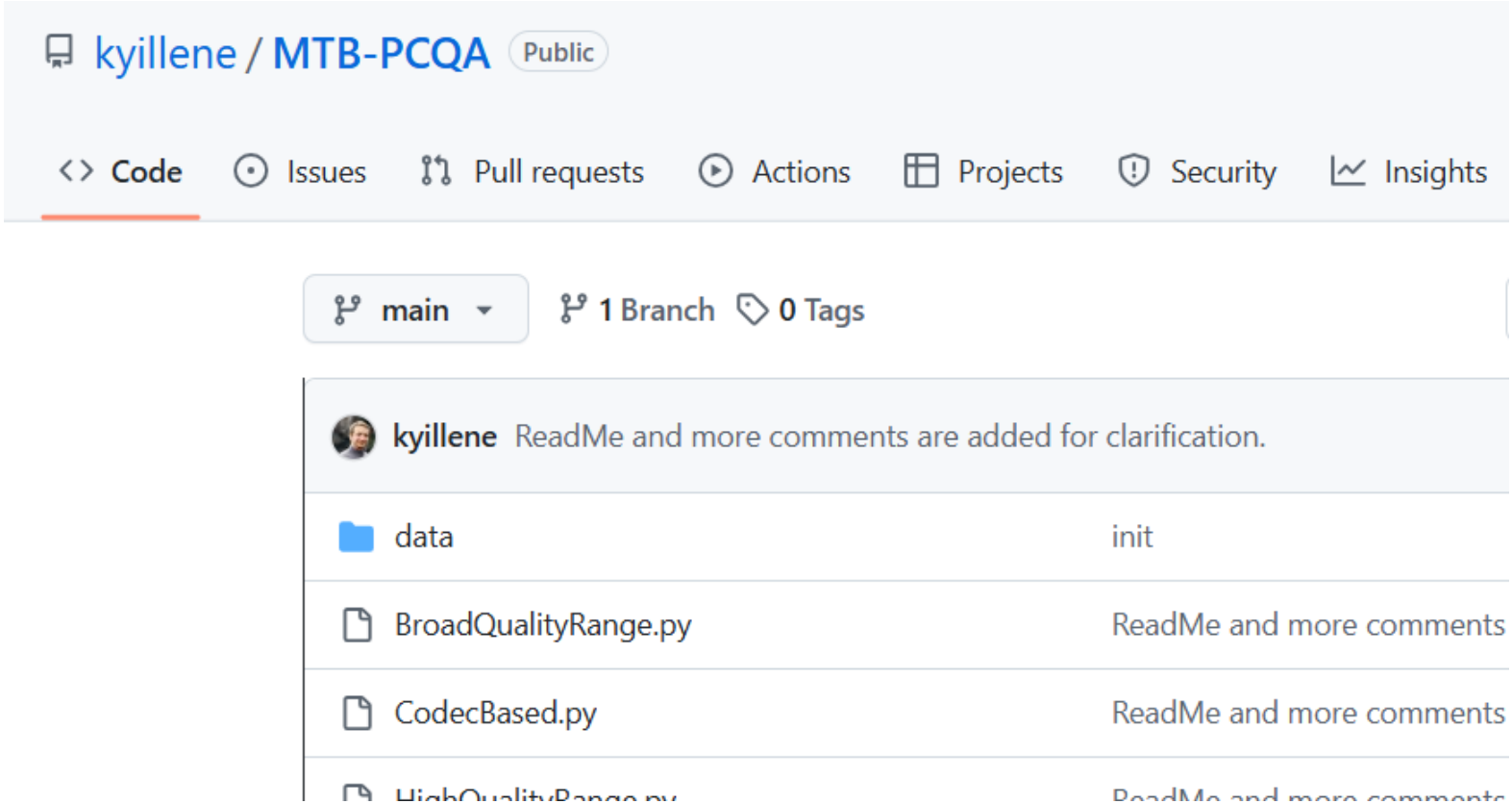
Track-based Evaluation Methodology



Track-based eval.

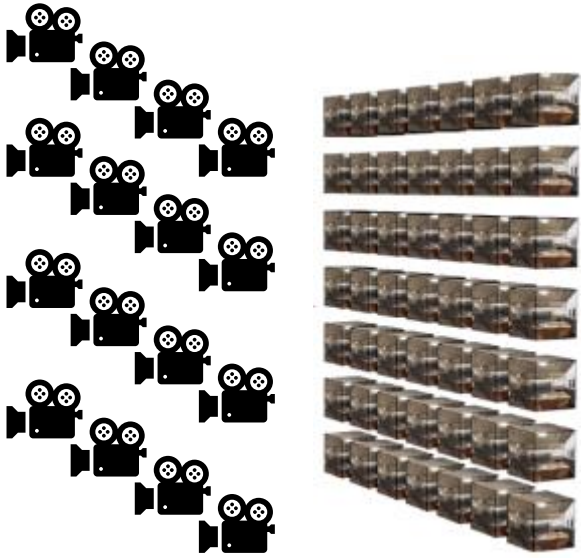
Ak, A., Zerman, E., Quach, M., Chetouani, A., Valenzise, G., & Le Callet, P. (2024, October). **A toolkit to benchmark point cloud quality metrics with multi-track evaluation criteria**. In *2024 IEEE International Conference on Image Processing (ICIP 2024)*.

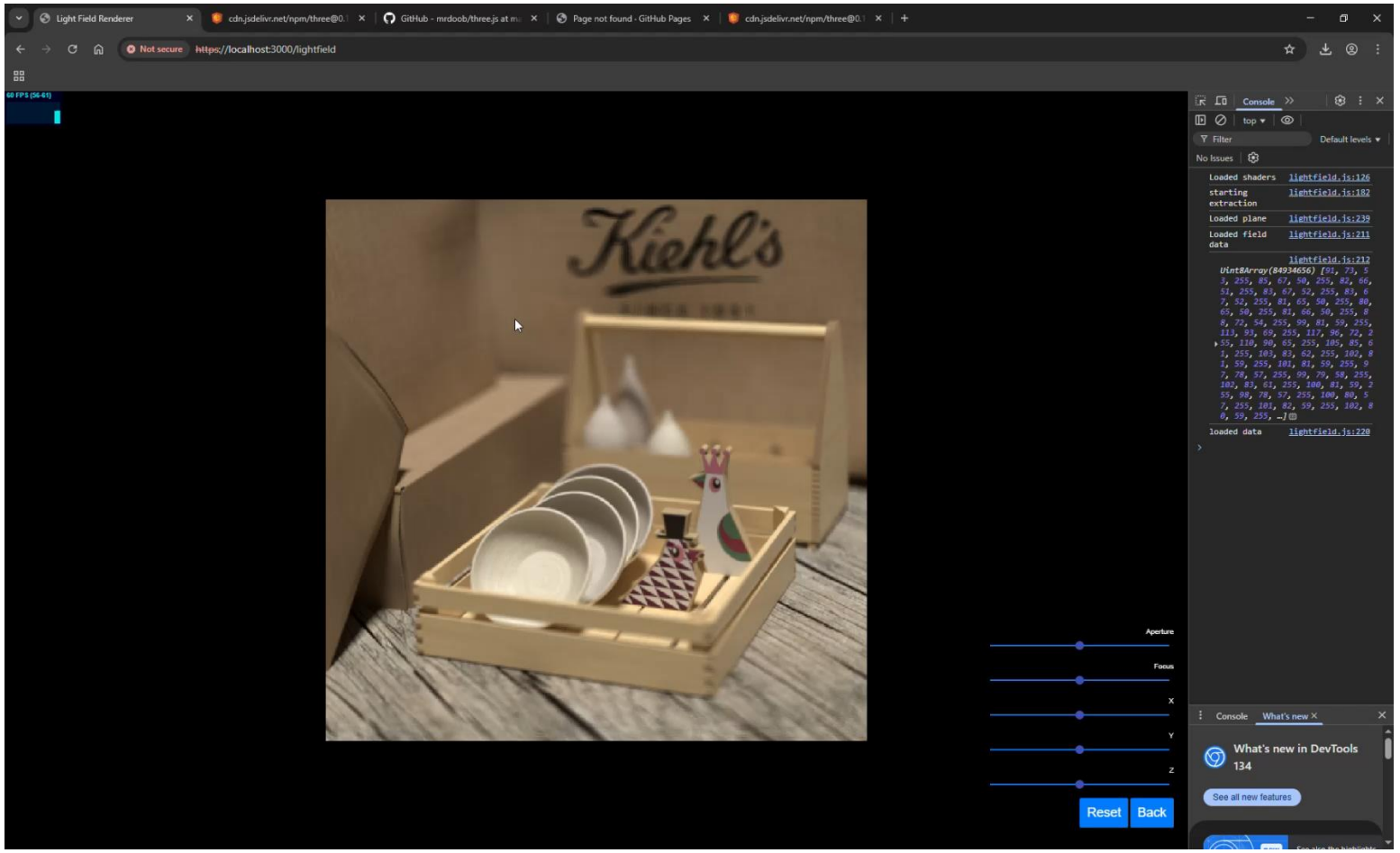
Track-based Evaluation Toolkit on GitHub



https://github.com/kyillene/MTB-PCQA

Light Fields for Telepresence





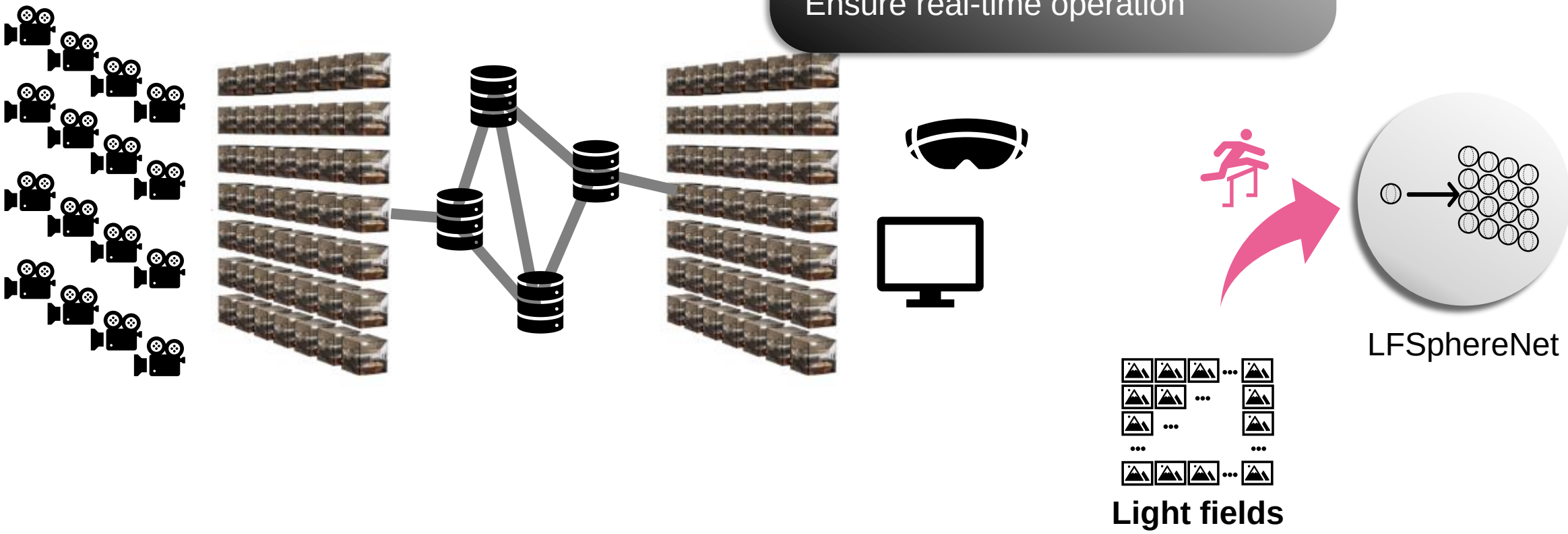
Challenge #3

Minimize the bandwidth load for telepresence applications

Motivation

Develop a method to reconstruct a LF from a single view

Ensure real-time operation



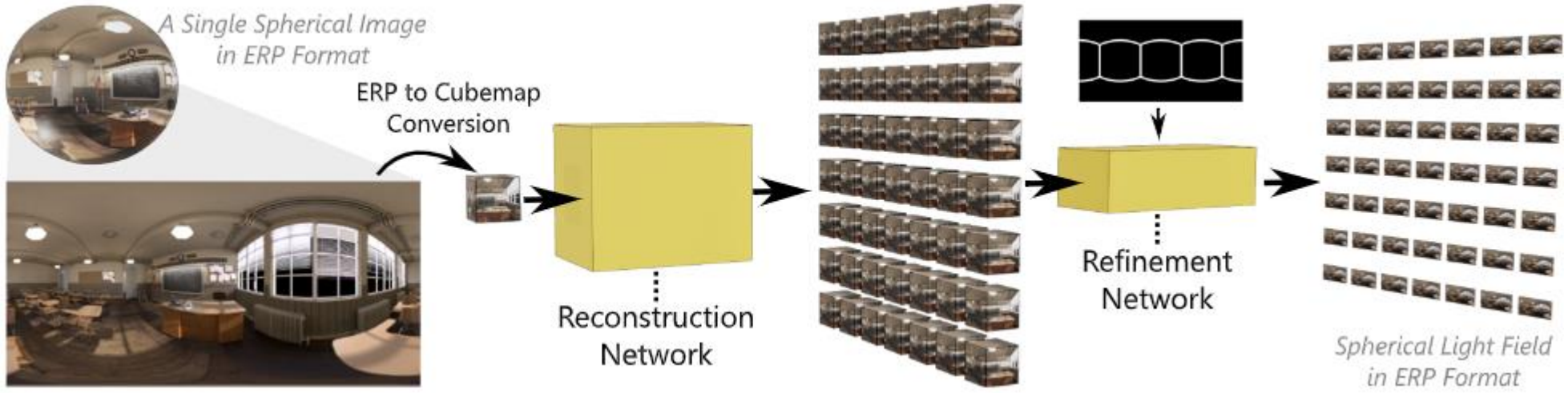
Challenge #3

Minimize the bandwidth load for telepresence applications

Mittuniversitetet
MID SWEDEN UNIVERSITY

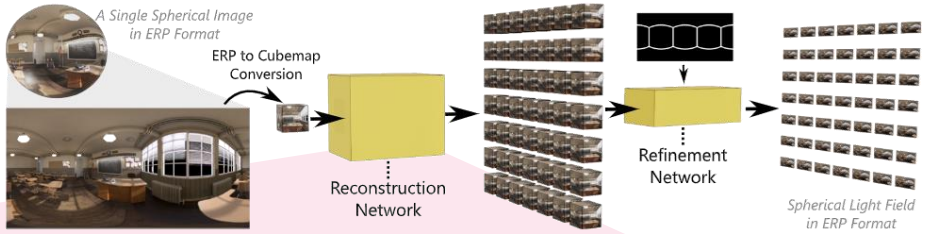
LFSphereNet

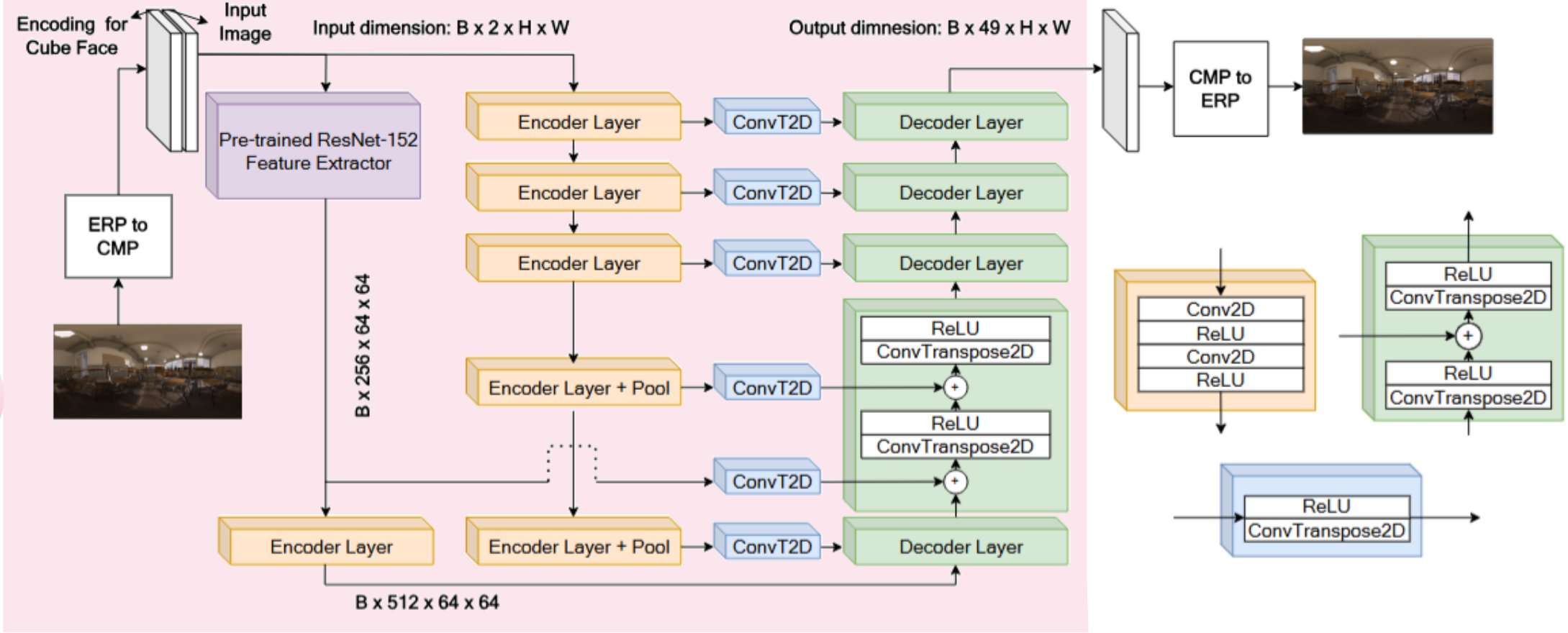
- A neural-network model to reconstruct a 7x7 spherical light field
- Transmission of only a single viewpoint



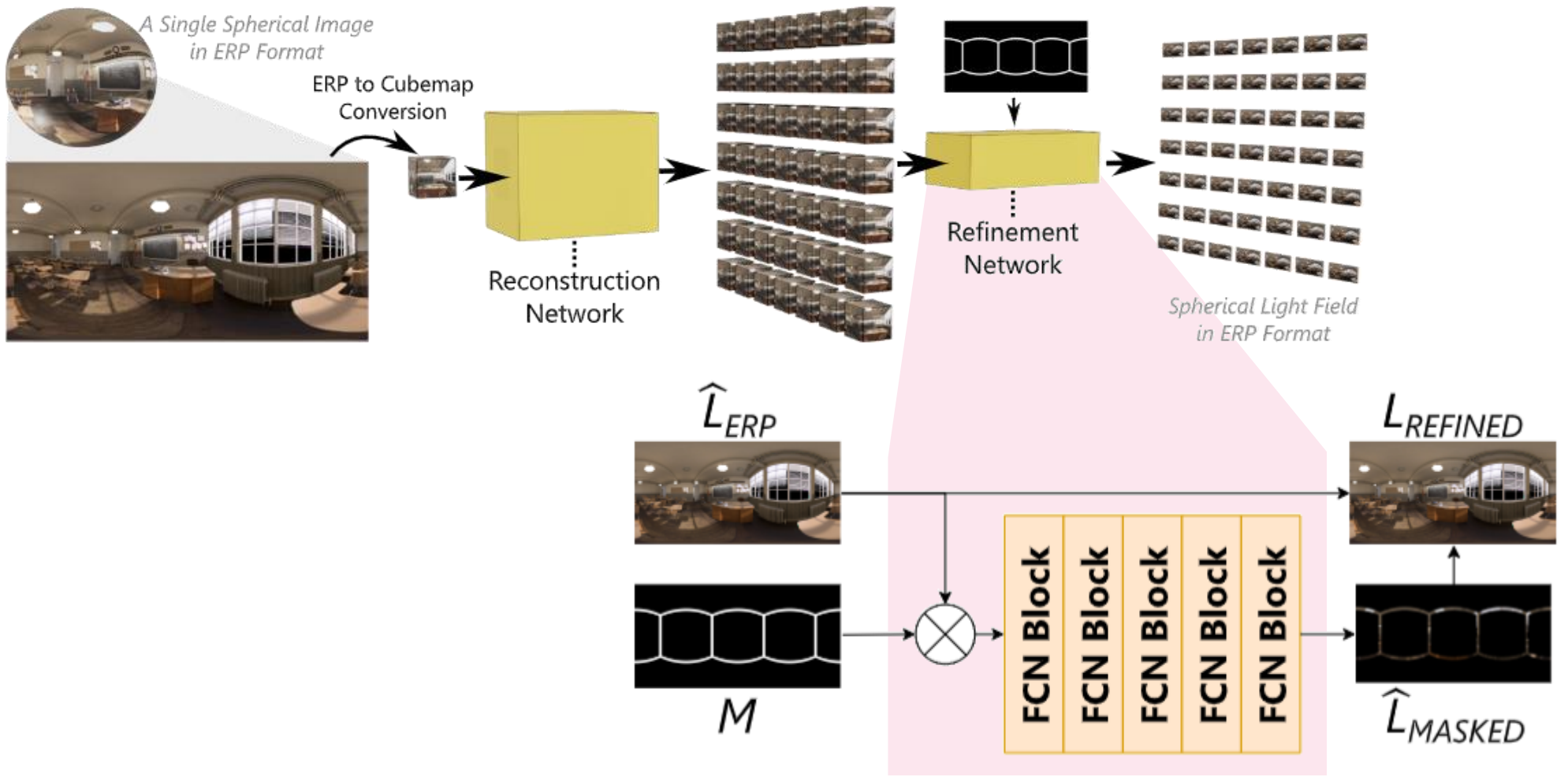
Gond, M., Zerman, E., Knorr, S., & Sjöström, M. (2023, November). **LFSphereNet: Real Time Spherical Light Field Reconstruction from a Single Omnidirectional Image**. In *Proceedings of the 20th ACM SIGGRAPH European Conference on Visual Media Production* (pp. 1-10).

LFSphereNet





LFSphereNet

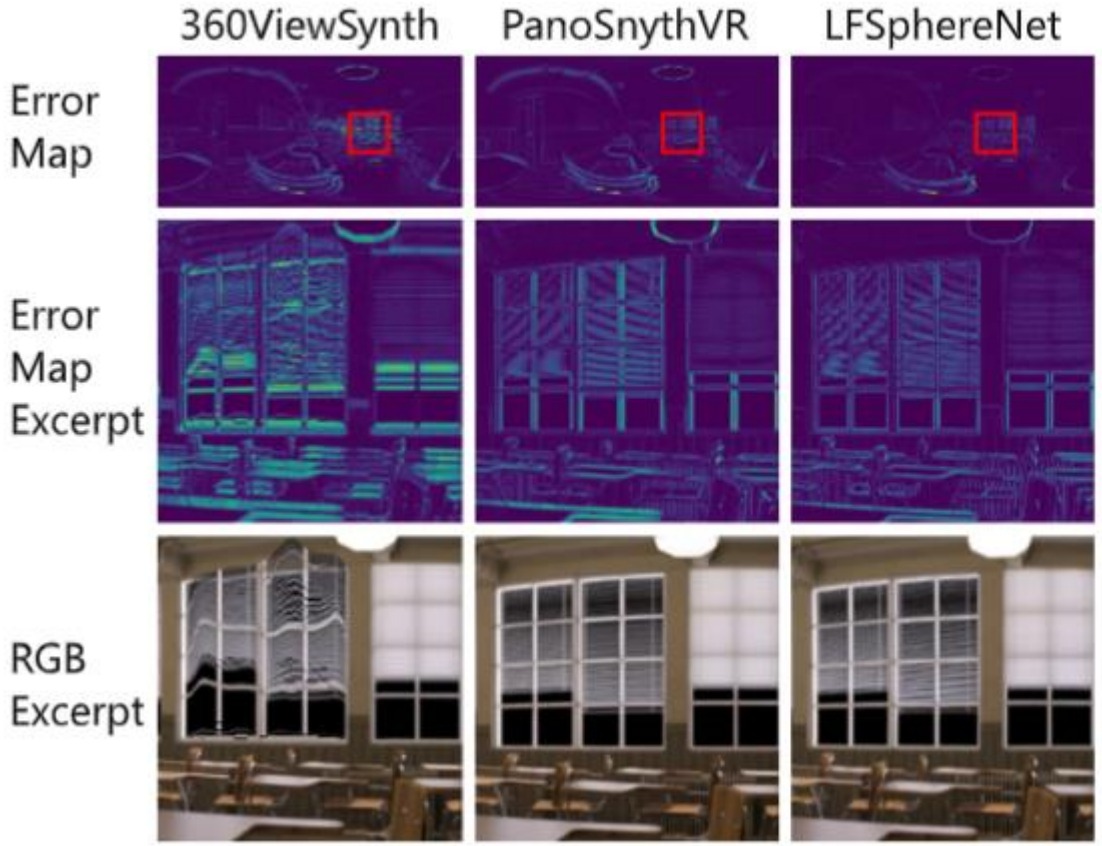


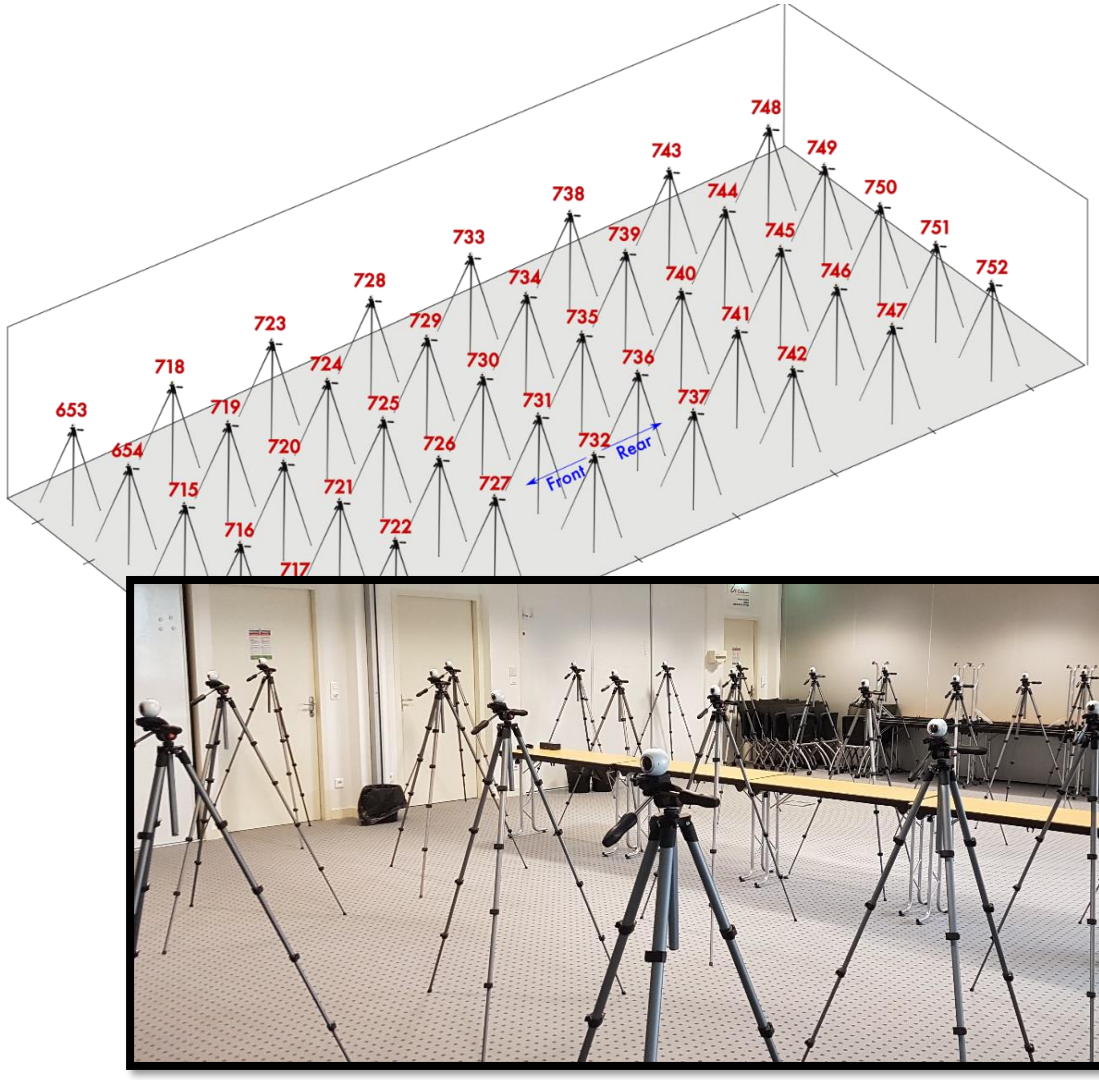
LFSphereNet evaluation

Metrics	LF Reconstruction Methods			
	[58]	[42]	LFSphereNet*	LFSphereNet
DISTS↓	0.1215	**0.0531**	0.0877	0.0880
LPIPS↓	0.1935	**0.0670**	0.0860	0.0825
PSNR↑	32.89	34.76	37.14	**37.45**
SSIM↑	0.6495	0.7878	0.9052	**0.9121**
Runtime↓	12.4845	2.7077	**0.0589**	0.0606

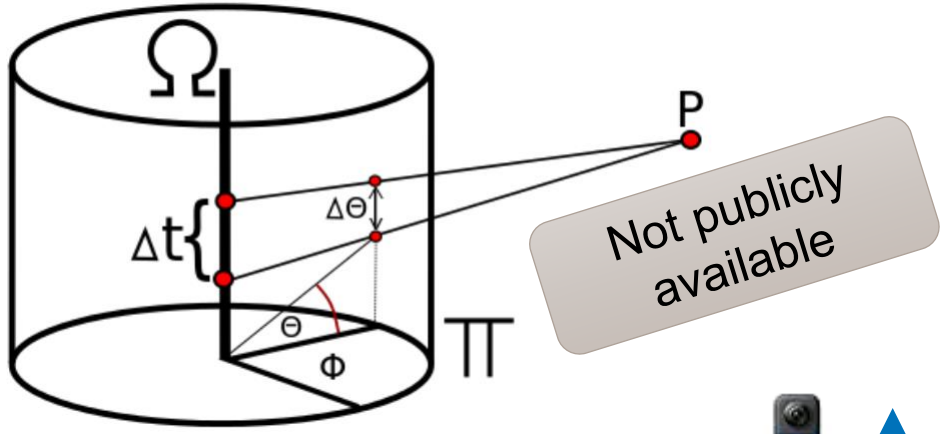
LFSphereNet* is without refinement network

- Uses a computer-generated (Blender) database for training and testing
- What about real-world datasets?
  - Larger baseline





Maugey, T., Guillo, L., & Cam, C. L. (2019, June). **FTV360: A multiview 360 video dataset with calibration parameters**. In *Proceedings of the 10th ACM Multimedia Systems Conference* (pp. 291-295).



Not publicly available

Krolla, B., Diebold, M., Goldlücke, B., & Stricker, D. (2014, September). **Spherical Light Fields**. In *BMVC* (No. 67.1–67.12).

Motivation

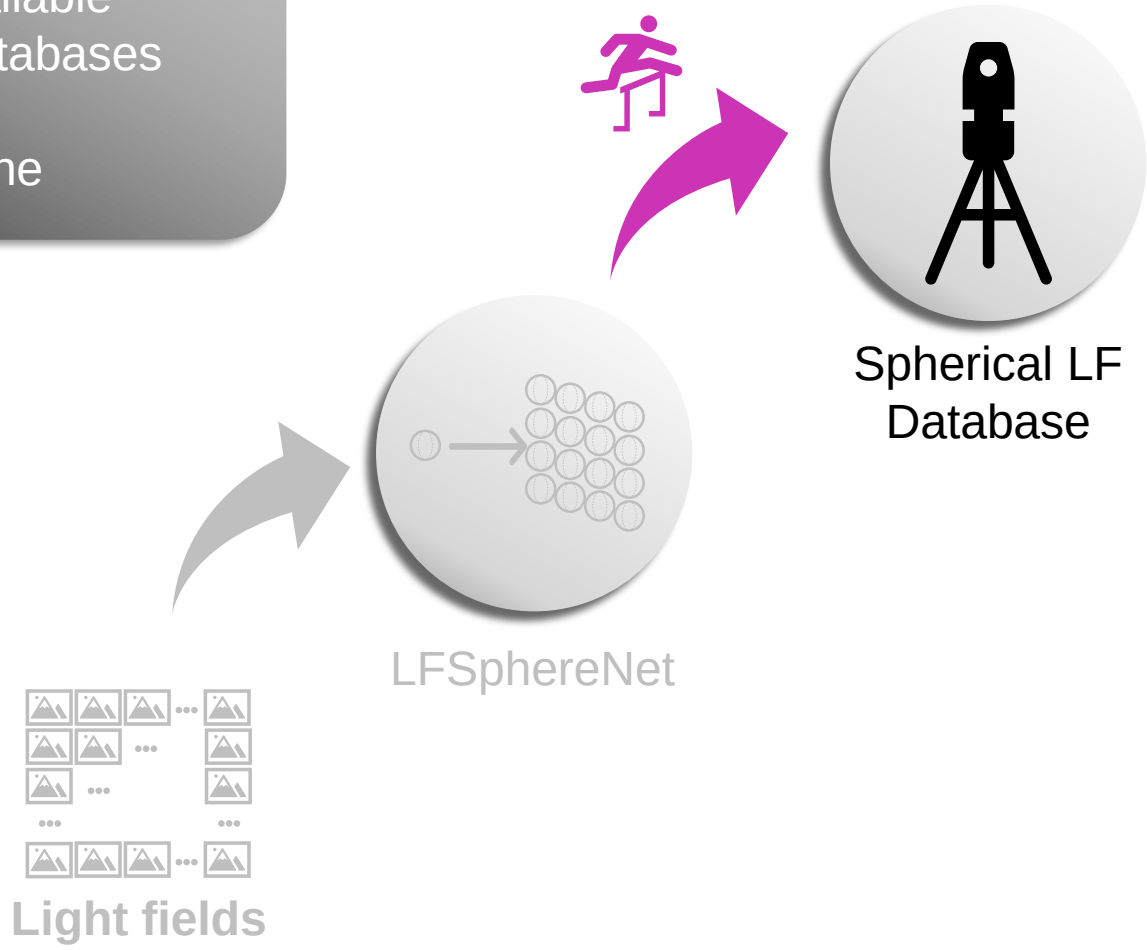
No suitable publicly available spherical light field databases

A need for larger baseline

Challenge #4

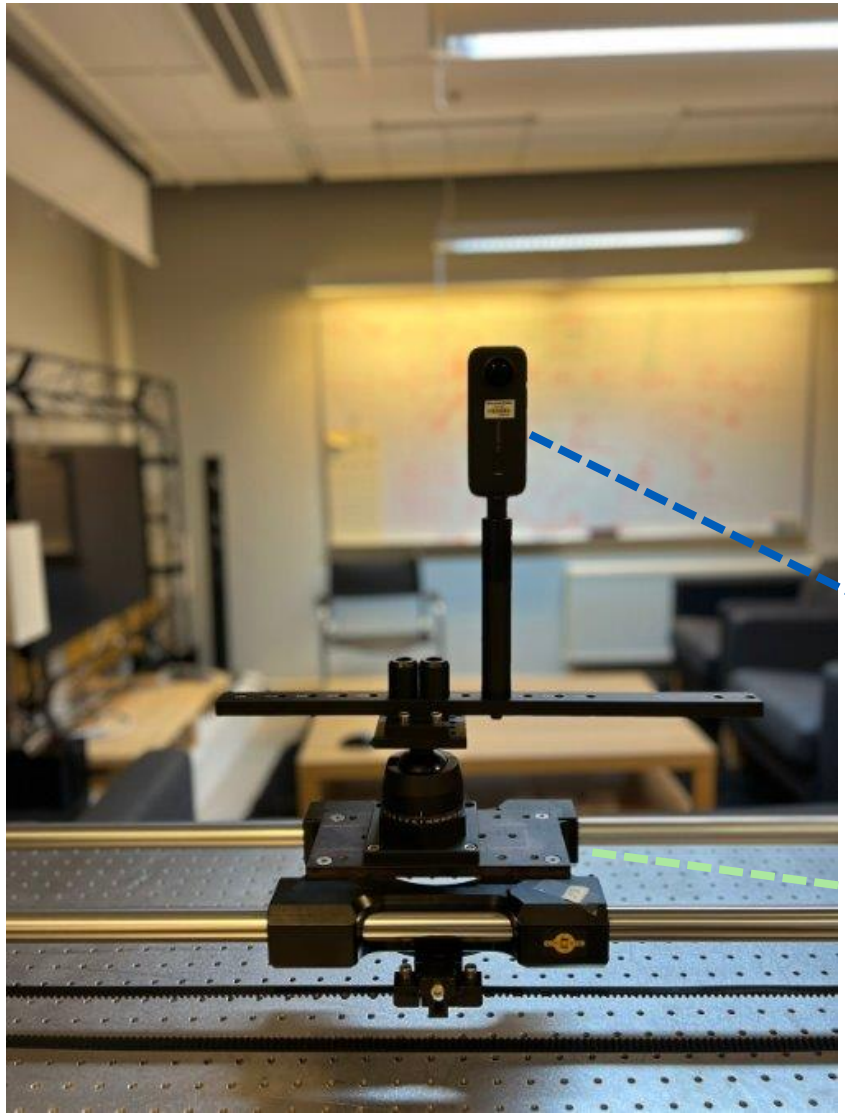
Positioning the camera

Lighting and color balance issues



Spherical LF Capture - Setup

Spherical LF Database

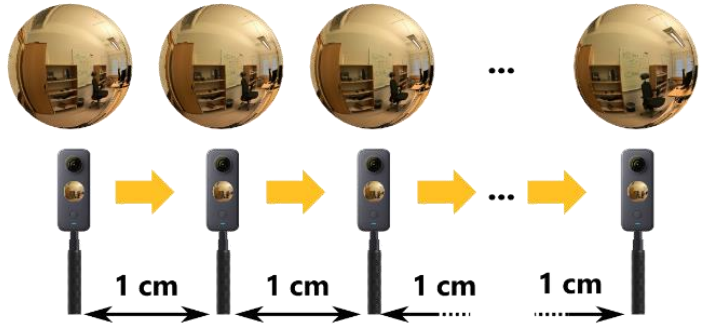


Challenge #4

Positioning the camera

Lighting and color balance issues

Insta 360 x3
Omnidirectional camera



Cinevate Atlas 200
Dolly

To ensure consistent and accurate movement

Challenges encountered in the pilot

- Positioning & static vs. Dynamic
- Lighting

Spherical LF Database

Challenges encountered in the pilot

- Positioning & static vs. Dynamic
- Lighting – (Clouds, passing objects, etc.)

Spherical LF Capture - Data




. . .


1 # 2 # 3 # 4 # 59 # 60

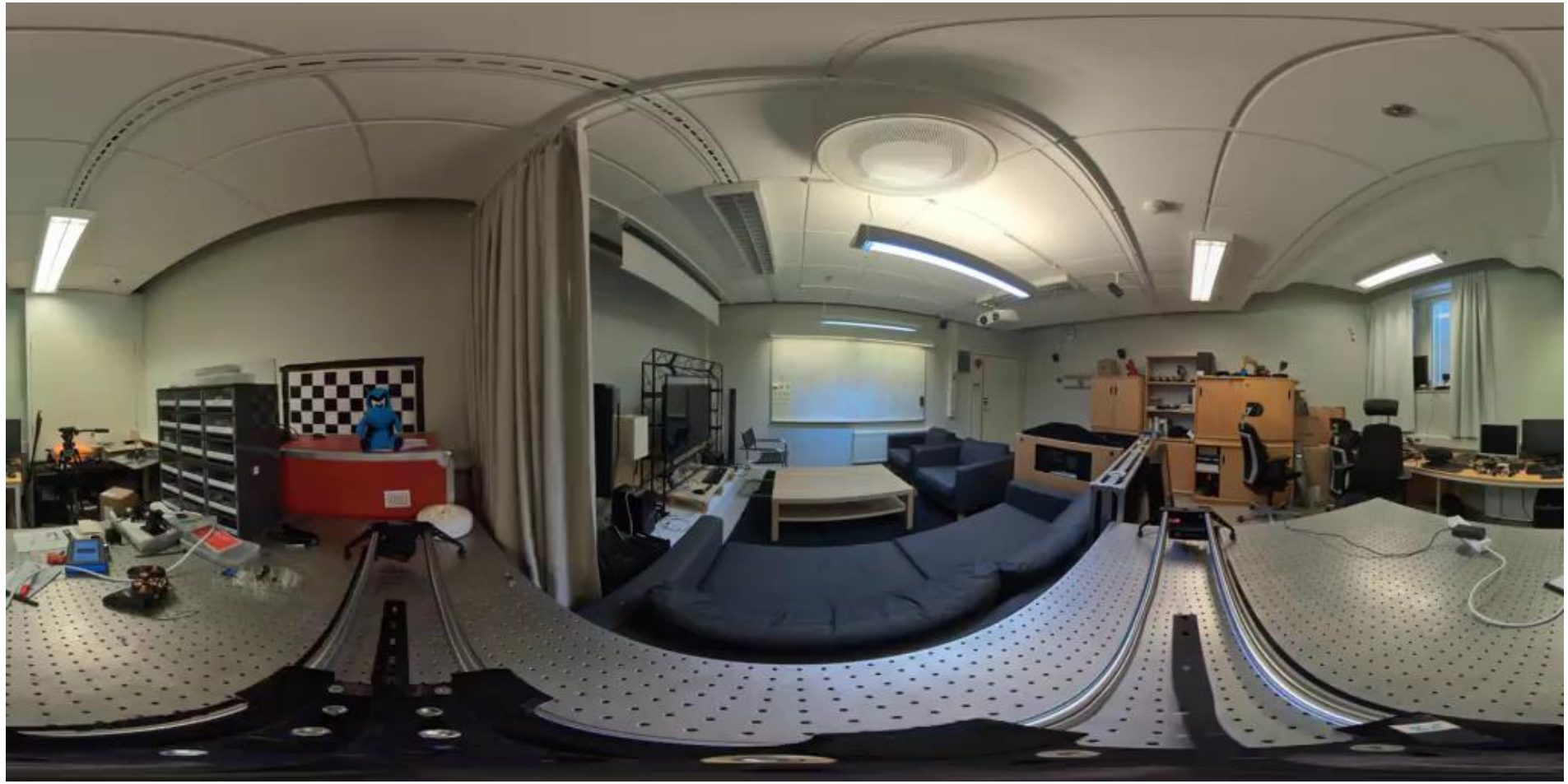
Spherical LF Database

1 x 60 spherical light field
ERP - 5120 × 2560

20 different indoor scenes

- Mid Sweden University, Sundsvall Campus

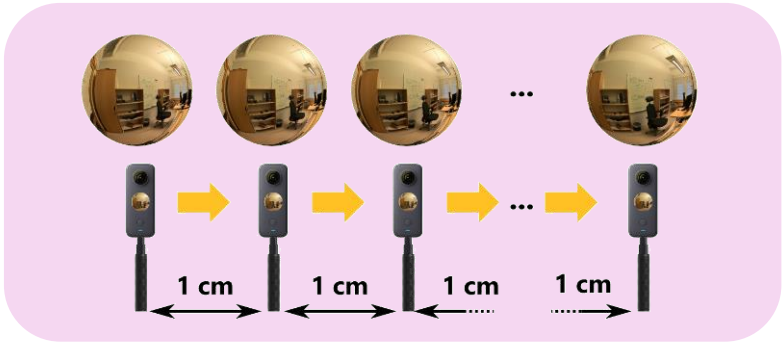
Spherical LF Capture - Data



Two Alternatives for Telepresence Applications

Spherical LF Database

Take all views



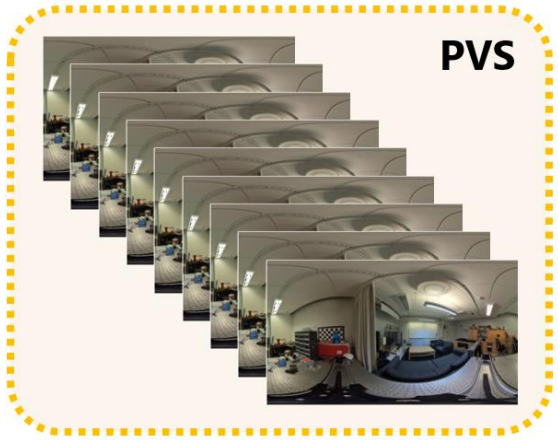
Take only center view

Compression

Decompression

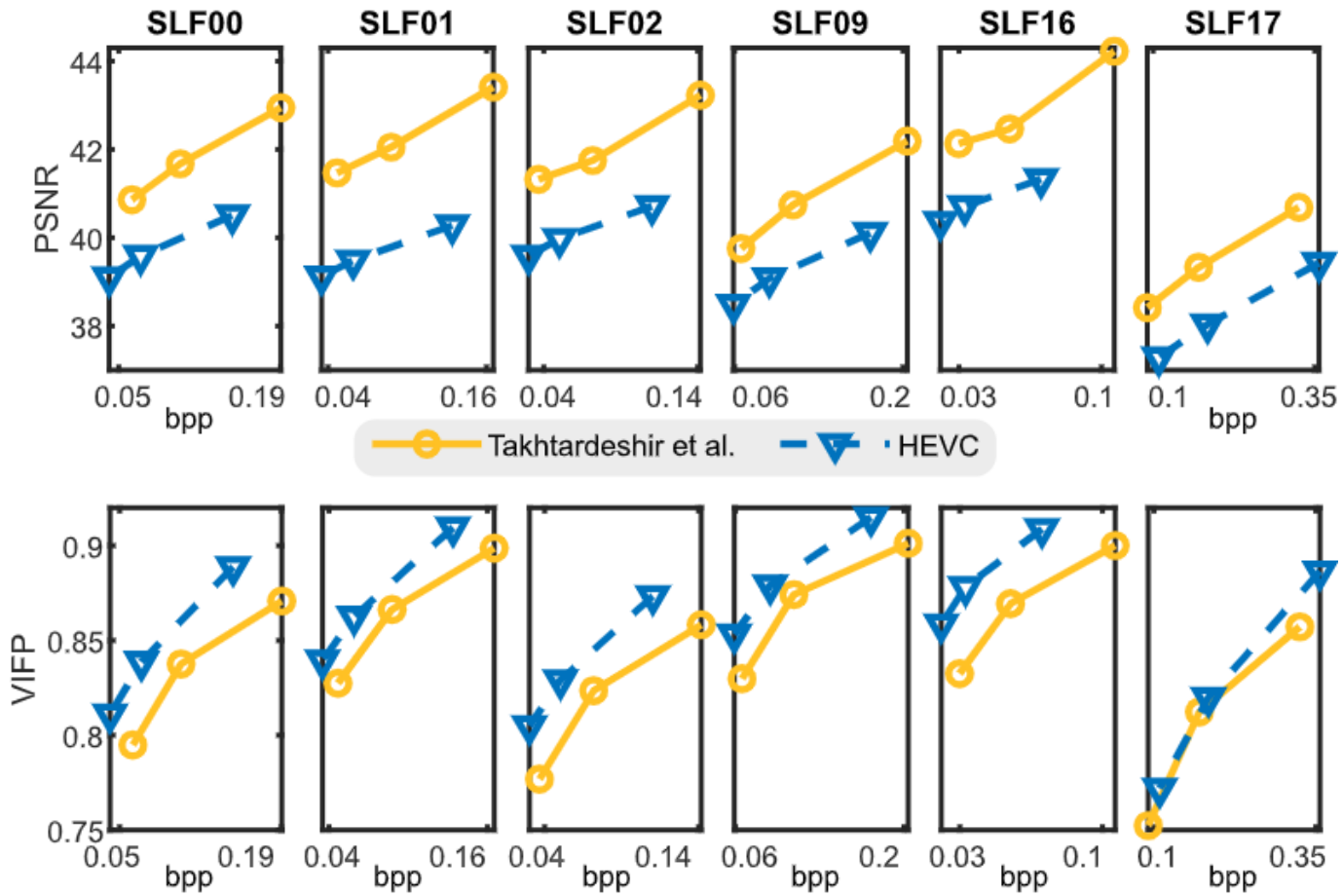
View Synthesis

SLFDB Compression



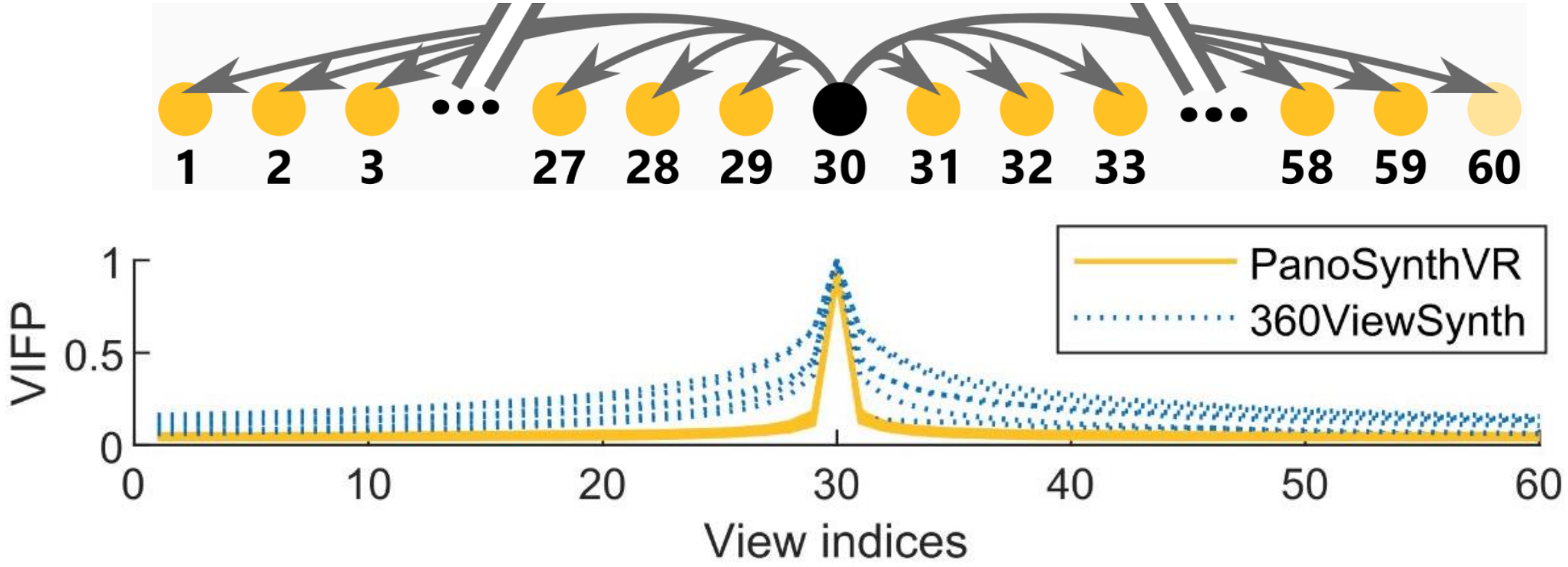
Compression

Bitstream



SLFDB View Synthesis

Spherical LF Database

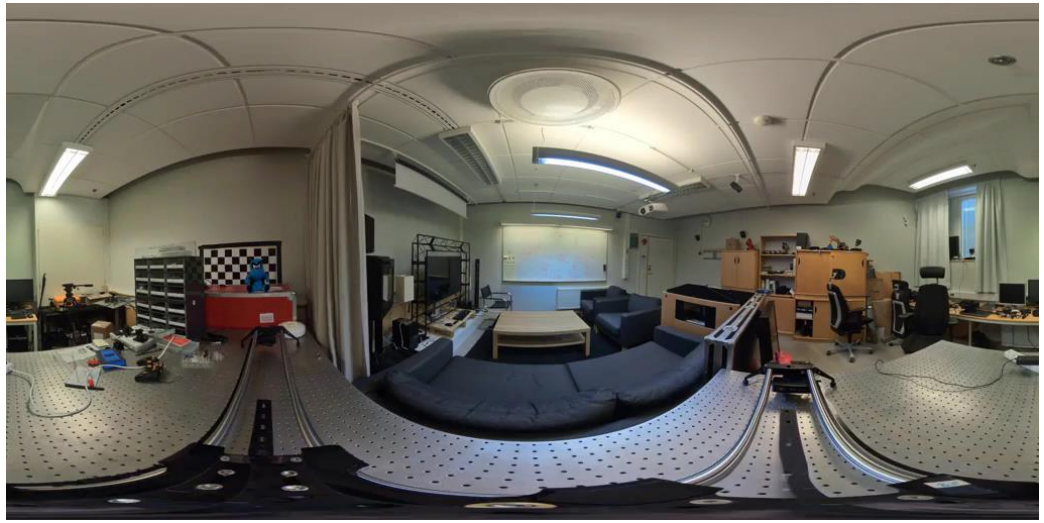


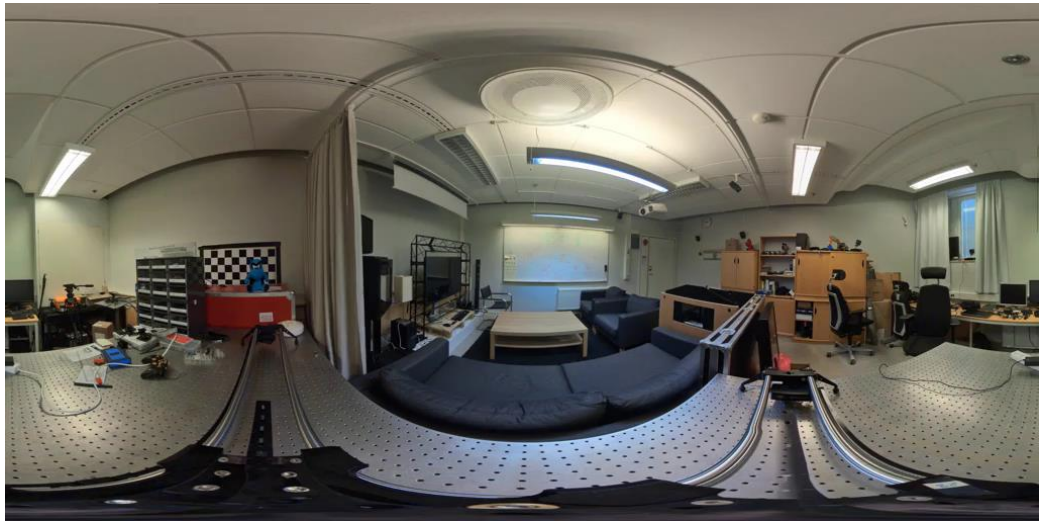
SLFDB View Synthesis – small baseline

Mittuniversitetet
MID SWEDEN UNIVERSITY

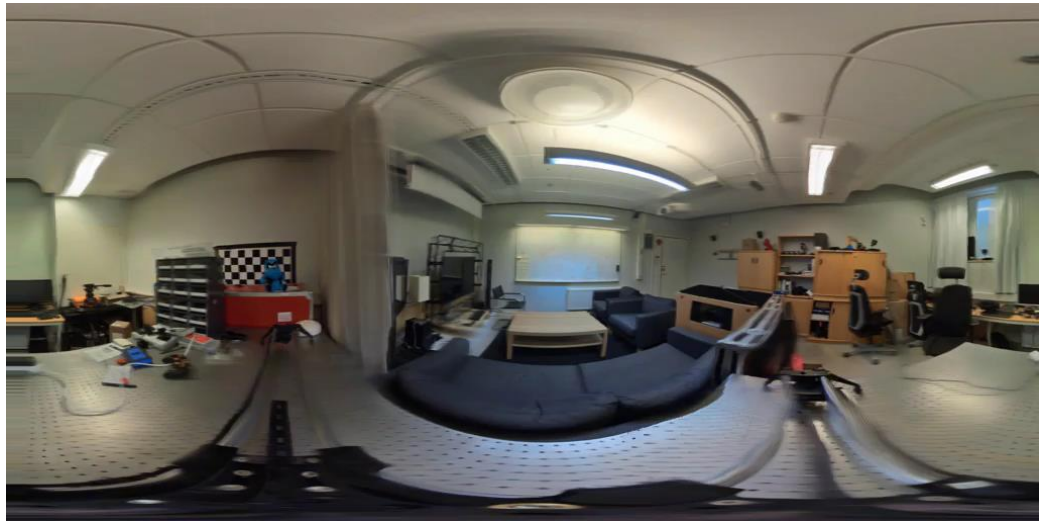
Spherical LF Database

Ground Truth





360ViewSynth



PanoSynthVR

SLFDB View Synthesis – small baseline

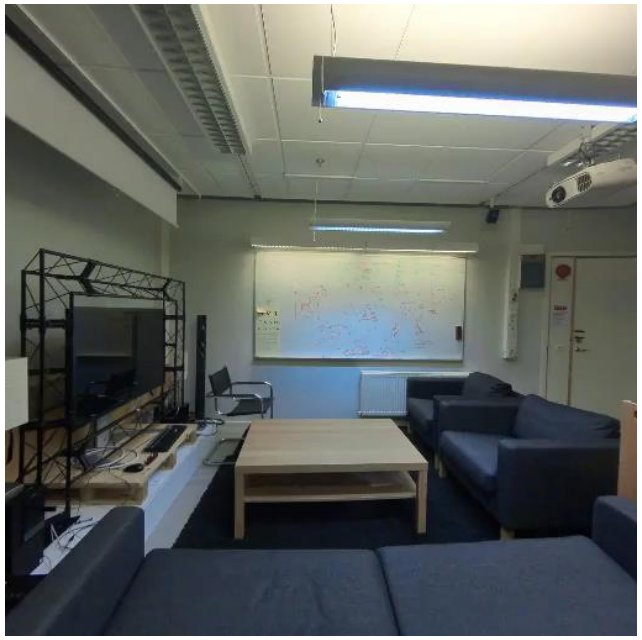
Mittuniversitetet
MID SWEDEN UNIVERSITY

Spherical LF Database

360ViewSynth	Ground Truth	PanoSynthVR

SLFDB View Synthesis – small baseline

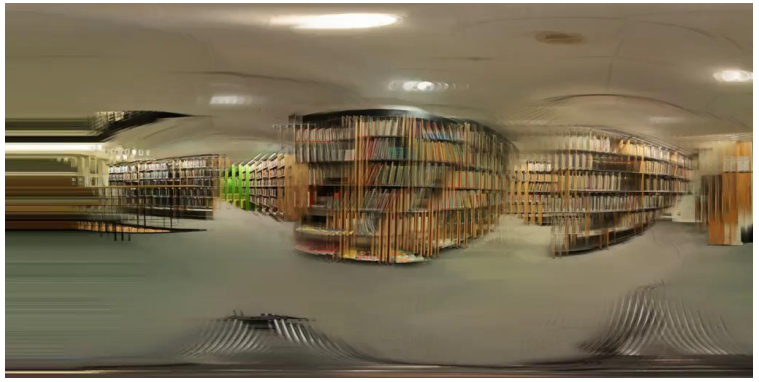
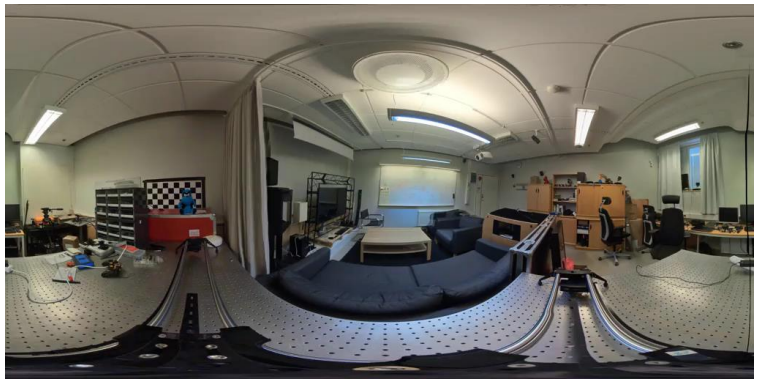
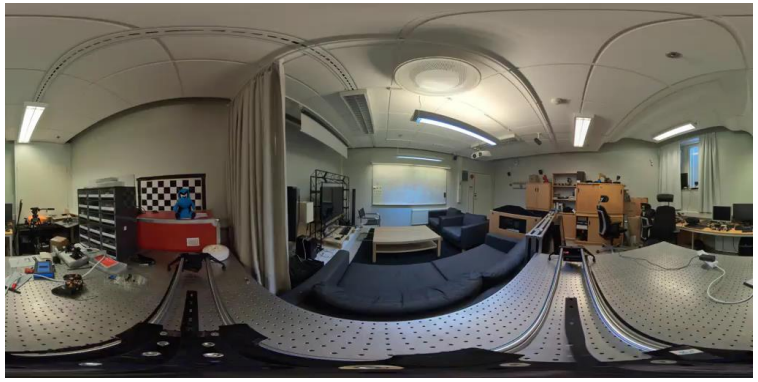
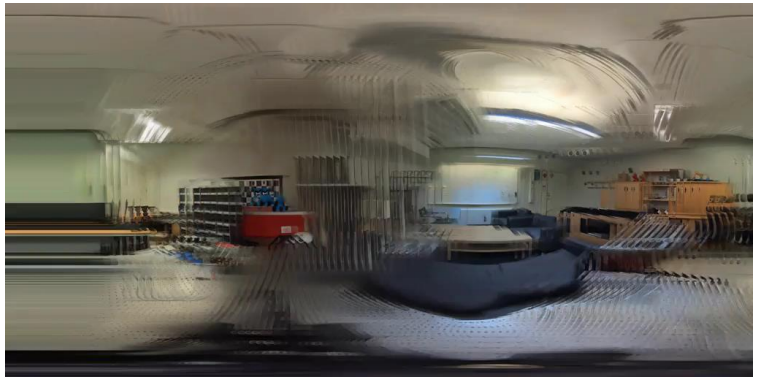
Spherical LF Database

360ViewSynth	Ground Truth	PanoSynthVR
		

SLFDB View Synthesis – large baseline

Mittuniversitetet
MID SWEDEN UNIVERSITY

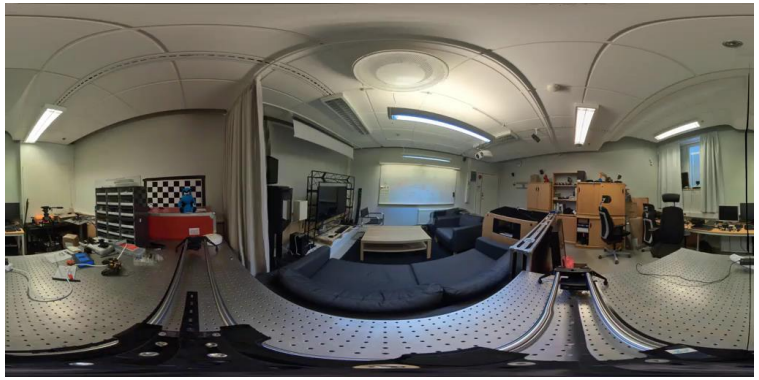
Spherical LF Database

360ViewSynth	Ground Truth	PanoSynthVR
		
		

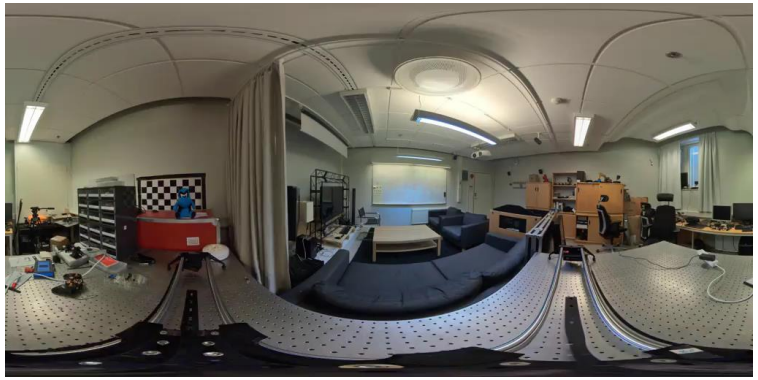
Open challenges

- Positioning & static vs. Dynamic
- Lighting – (Clouds, passing objects, etc.)
- View synthesis – large baseline

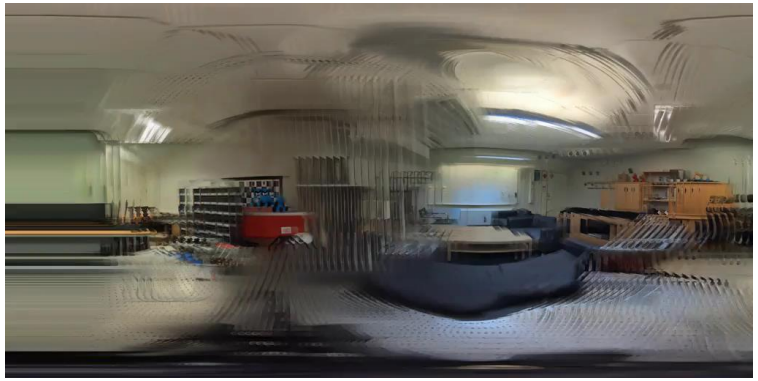
360ViewSynth



Ground Truth



PanoSynthVR



SLFDB

- Public spherical light field database
  - 1x60 views
- Validated on two use cases
  - Spherical LF compression
  - View synthesis
- Many challenges remain
  - Lighting, capture
  - Compression
  - View synthesis

Preview the database:

https://3zadessg.github.io/SLFDB/SLFDB.html

Zerman, E., Gond, M., Takhtardeshir, S., Olsson, R., & Sjöström, M. (2024, June). **A Spherical Light Field Database for Immersive Telecommunication and Telepresence Applications**. In *2024 16th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE.

Take-Away Messages

- Challenge-driven dataset creation is essential.
  - **We are at the frontier!**
- Ecological validity is key for metric evaluation.
  - Evaluation vs. real-world performance.
- Light field telepresence adds new constraints.
  - **Innovative reconstruction methods** and **realistic datasets** are needed.
- Each challenge shape better tools and datasets.
  - Tackling them moves the field closer to real-world Social XR applications.

Mittuniversitetet
MID SWEDEN UNIVERSITY

Dr. Emin Zerman

Associate Senior Lecturer

Mid Sweden University

Department of Computer and Electrical Engineering

emin.zerman@miun.se