Hologram generation via deep-learning

Tomoyoshi Shimobaba,
Fan Wang, Masato Shotoku, Kai Kumano, Yudai Fujima, Keisuke Ono,
Tomoyoshi Ito

Chiba University, Japan

shimobaba@faculty.chiba-u.jp

About our lab

- 3 professors
- 1 postdoc
- 35 students
    - 7 Ph.D students
    - 16 master students
    - 12 undergraduate students

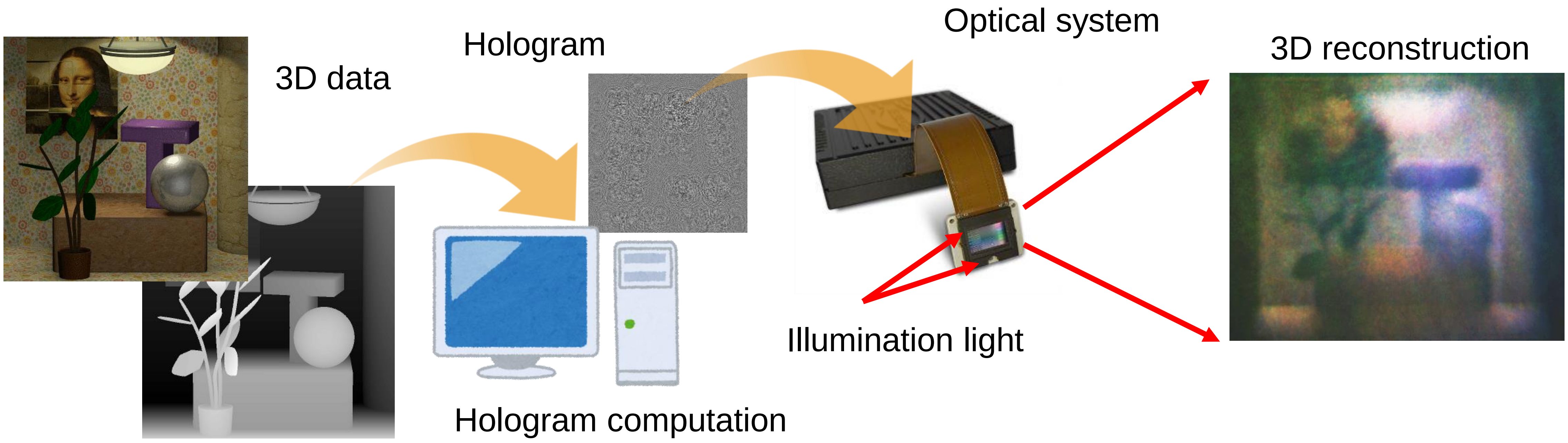
- Research topics
    - Holographic display and its related algorithms
    - Digital holography
    - Single-pixel imaging
    - Special-purpose computers using FPGA
    - Deep learning


Apr. 2025
Prof. Ito
Prof. Wang
Me
Dr. Starobrat
(WUT, Poland)

Holographic display

Holography can faithfully reproduce the wavefront of light

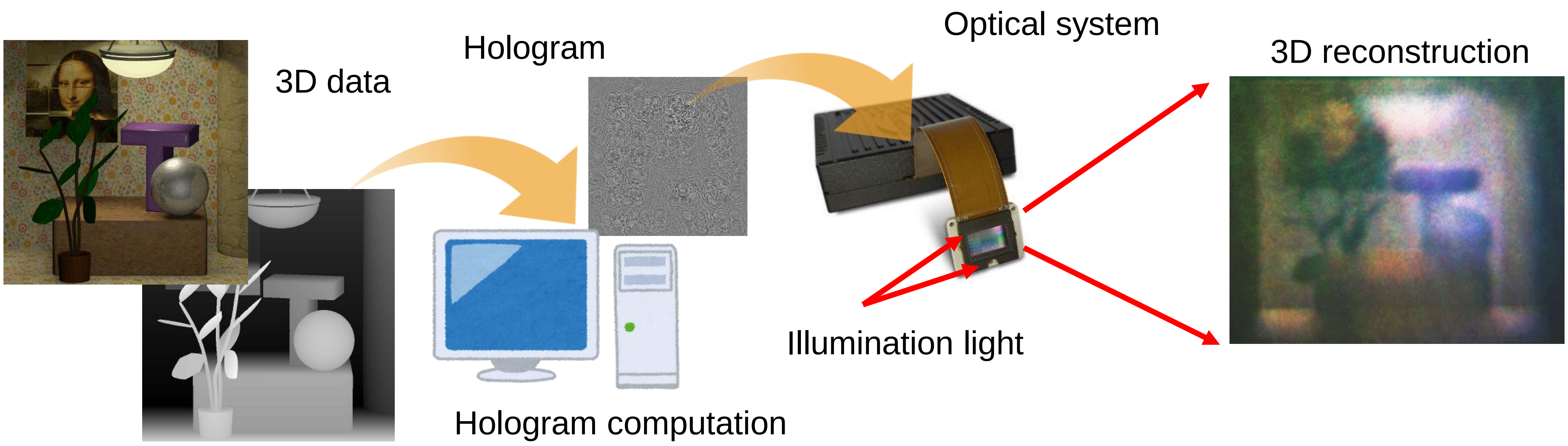
<u>The only method, the ultimate 3D technology</u>



Holographic display

Holography can faithfully reproduce the wavefront of light

<u>The only method, the ultimate 3D technology</u>

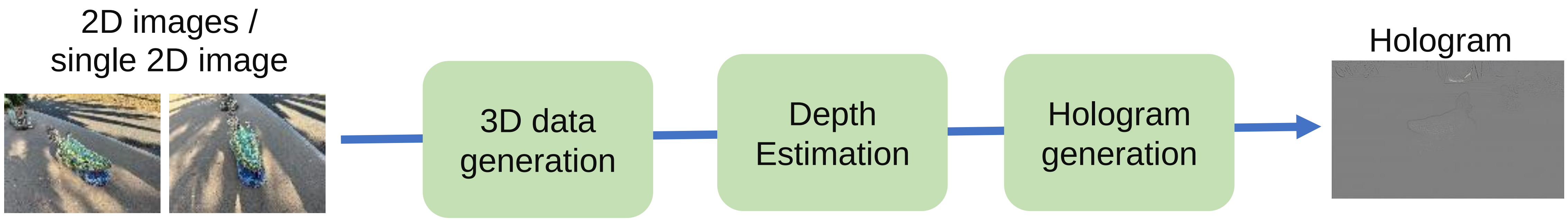


<u>What is difficult?</u>

- Hologram computation
- Cumbersome for 3D data generation
- Image quality
- Small image size and viewing angle

Topics

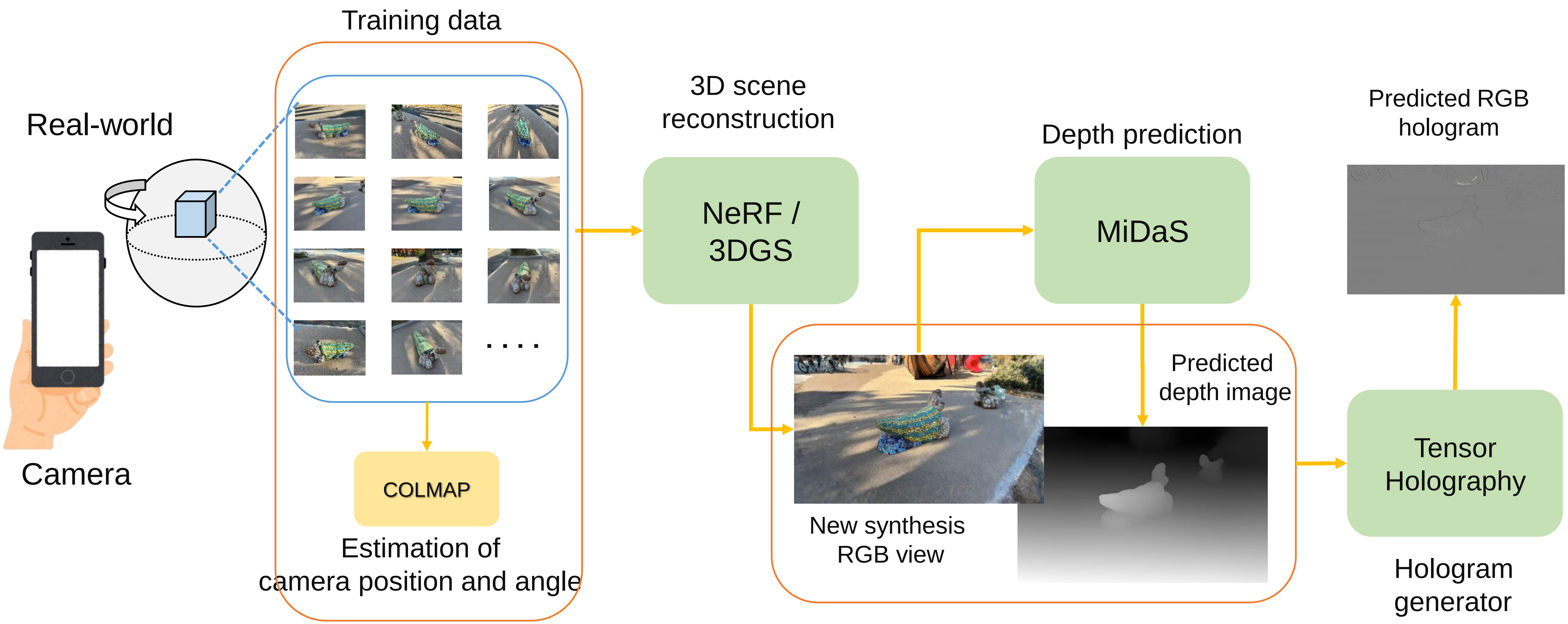
- Free viewpoint holographic displays using NeRF / 3DGS
- Free viewpoint holographic displays using Vision NeRF
- Binocular hologram generation
- 3D TV game using holographic signal converter



Topics

- Free viewpoint holographic displays using NeRF / 3DGS
- Free viewpoint holographic displays using Vision NeRF
- Binocular hologram generation
- 3D TV game using holographic signal converter

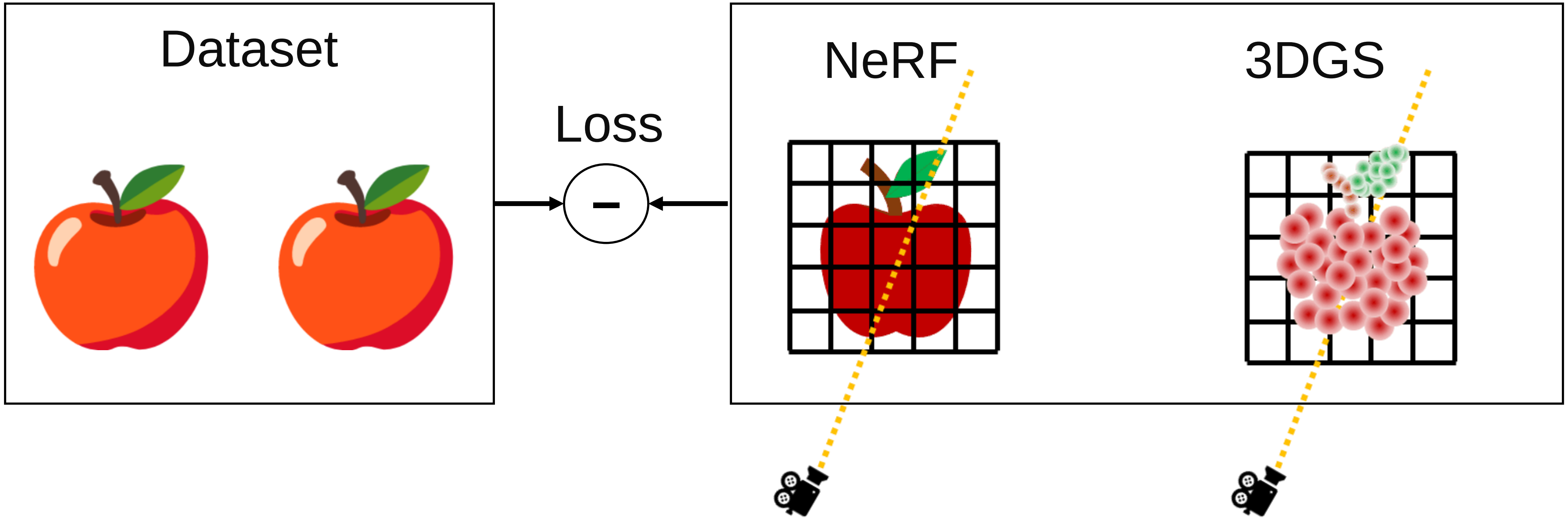
Free viewpoint holographic displays using NeRF / 3DGS



Kang, M., Wang, F., Kumano, K., Ito, T., & Shimobaba, T. (2024). Neural-radiance-fields-based holography. *Applied Optics*, *63*(28), G24-G29.

Kumano, K., Wang, F., Ono, K., Ito, T., & Shimobaba, T. Computer-Generated Hologram Based on 3d Gaussian Splatting. *Optica Engineering (accepted)*

Neural Radiance Field (NeRF) / 3D Gaussian Splatting (GS)



Representation of 3D scene

Neural network (Implicit)

Gaussian (Explicit)

Neural Radiance Field (NeRF) / 3D Gaussian Splatting (GS)

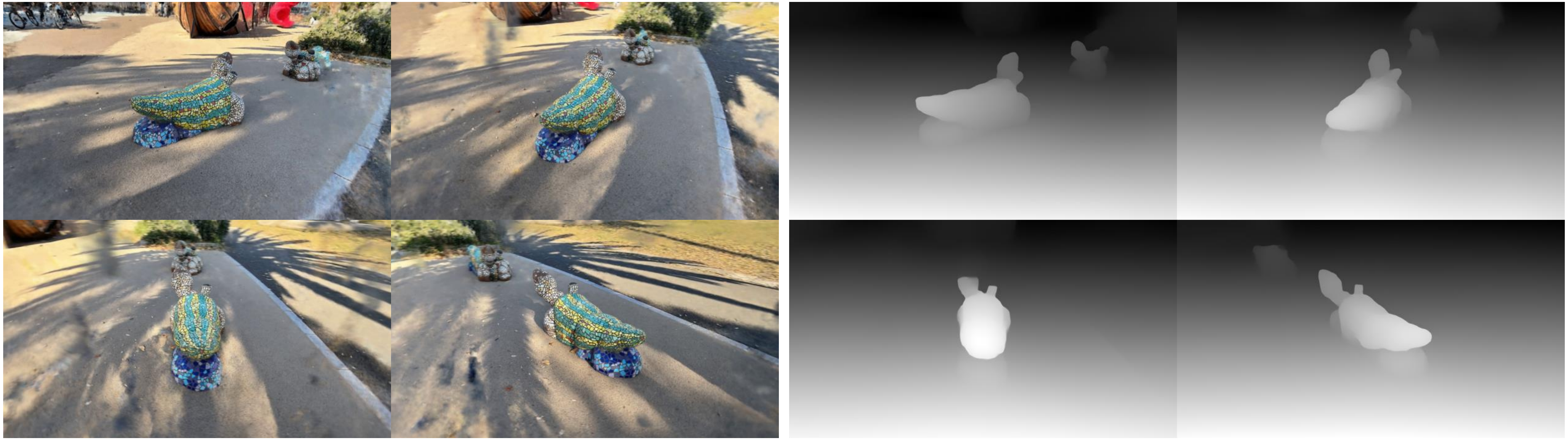
Dataset	NeRF	3DGS

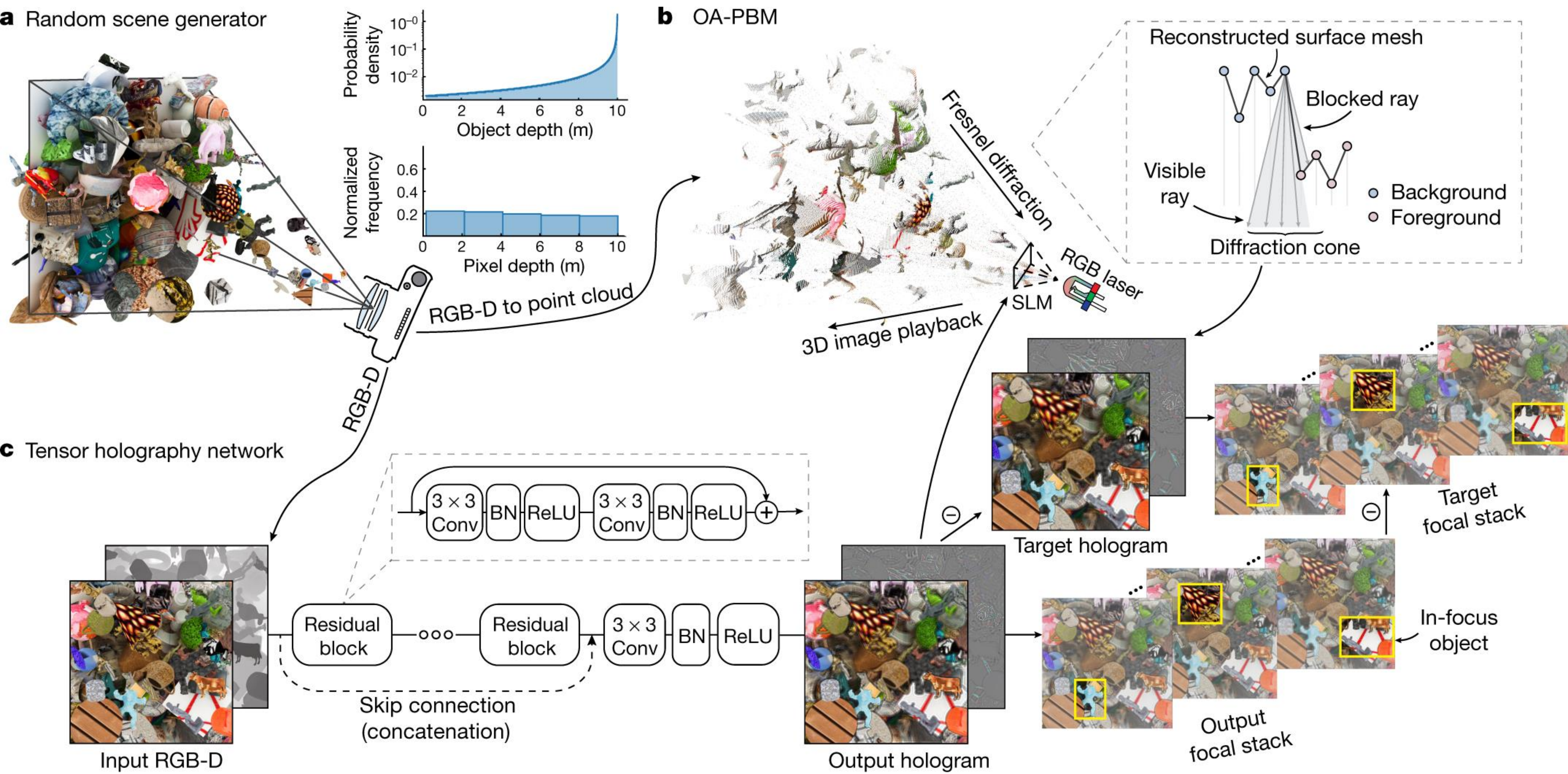
Representation of 3D scene	Neural network (Implicit)	Gaussian (Explicit)
Rendering speed	Slow	Fast

Depth prediction: MiDaS

Ranftl+ IEEE PAMI (2020)

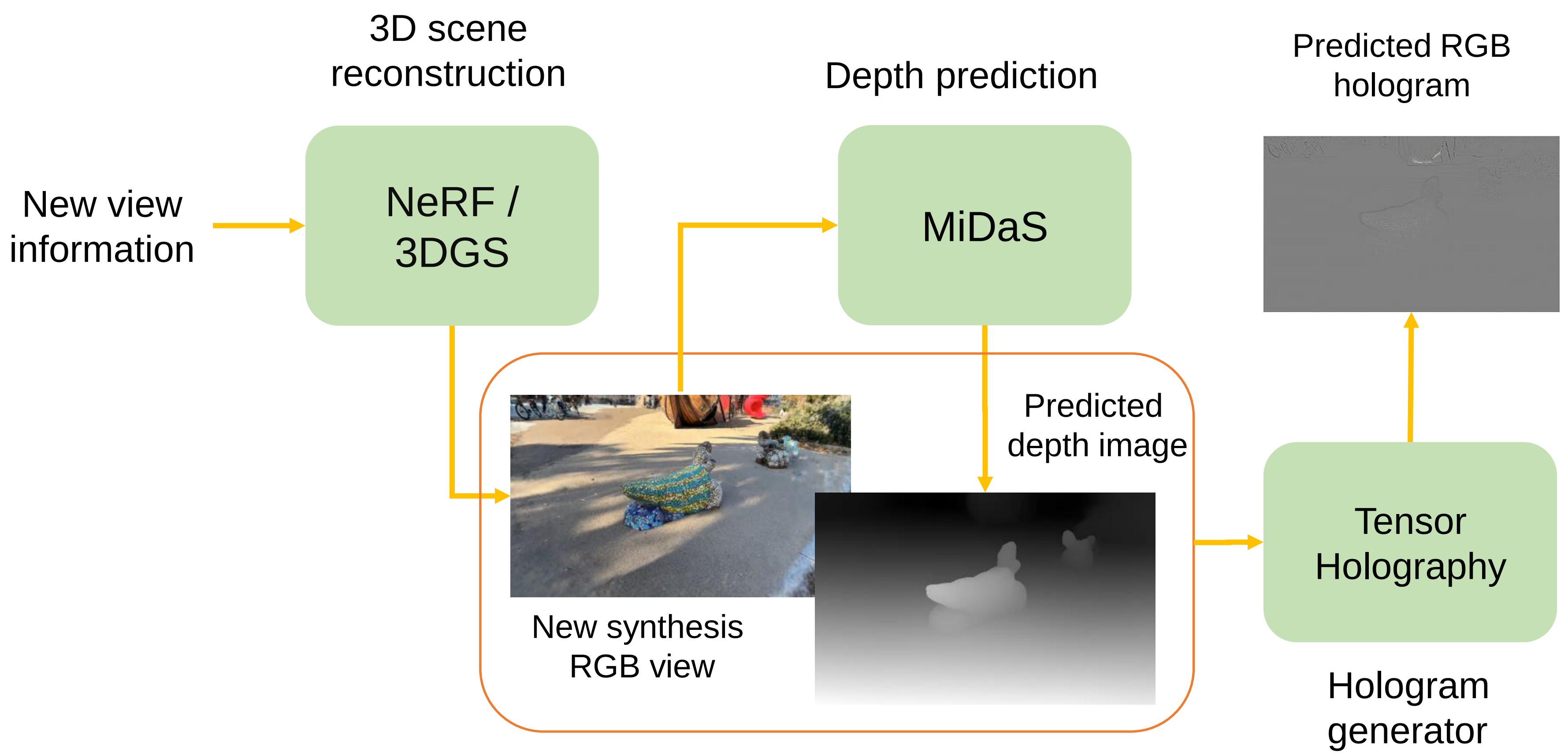
A monocular depth estimation model that can be used with zero-shot (no fine-tuning), trained on multiple diverse datasets



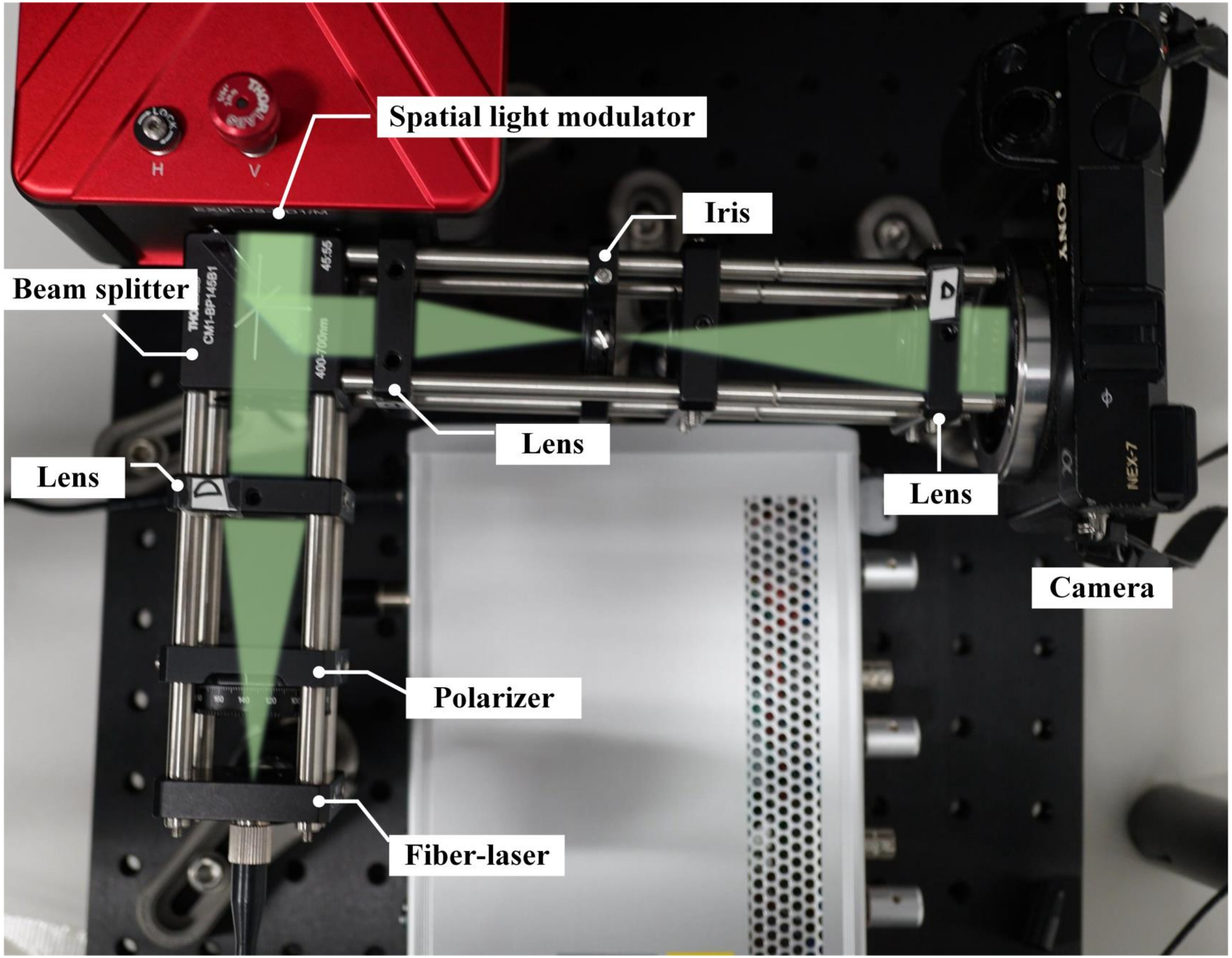


Shi+ Nature, 591 (2020)

Summary of the proposed pipeline

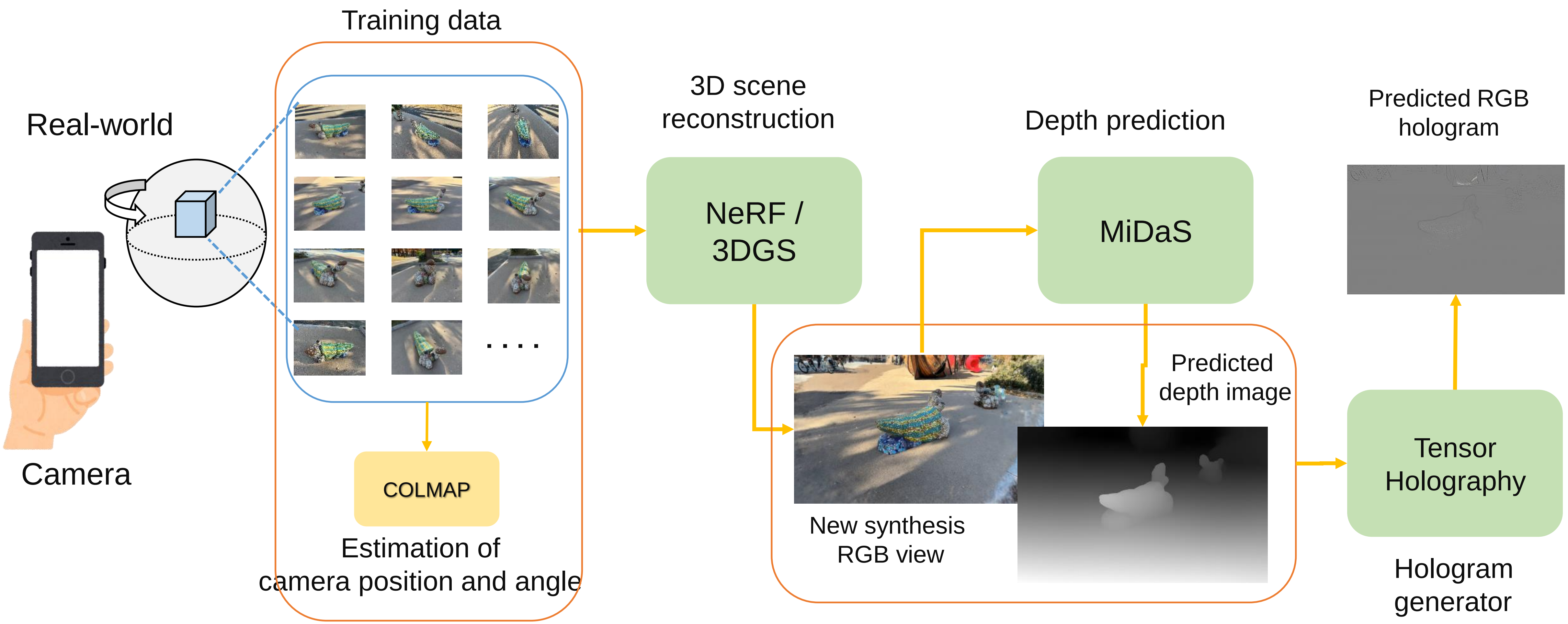


Optical experiments



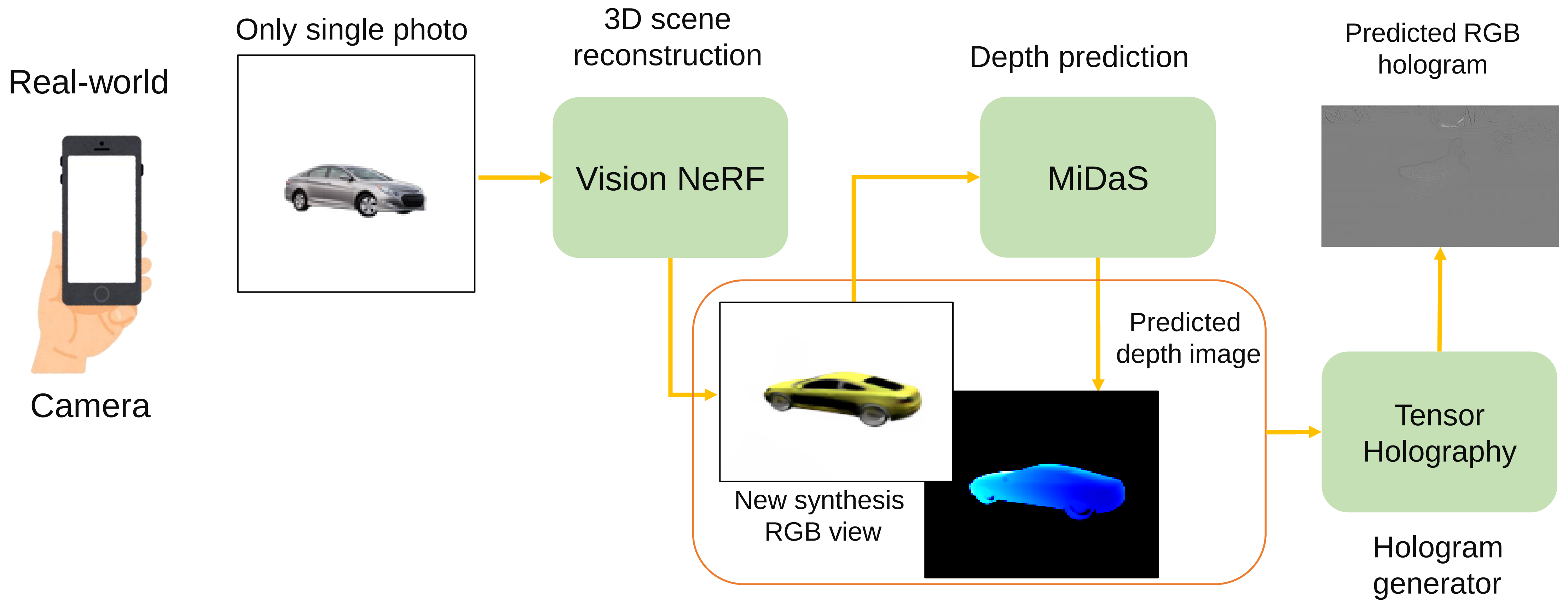
Prof. Fan Wang
Chiba Univ.

Problem of NeRF-based hologram generation



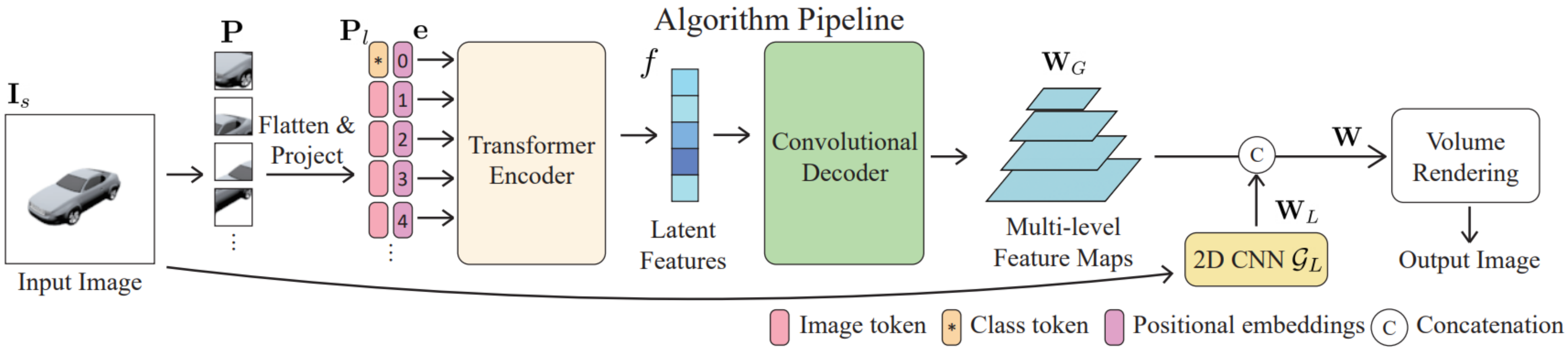
- Required many 2D images
- Camera pose estimation

Free viewpoint holographic displays using Vision NeRF



Vision NeRF

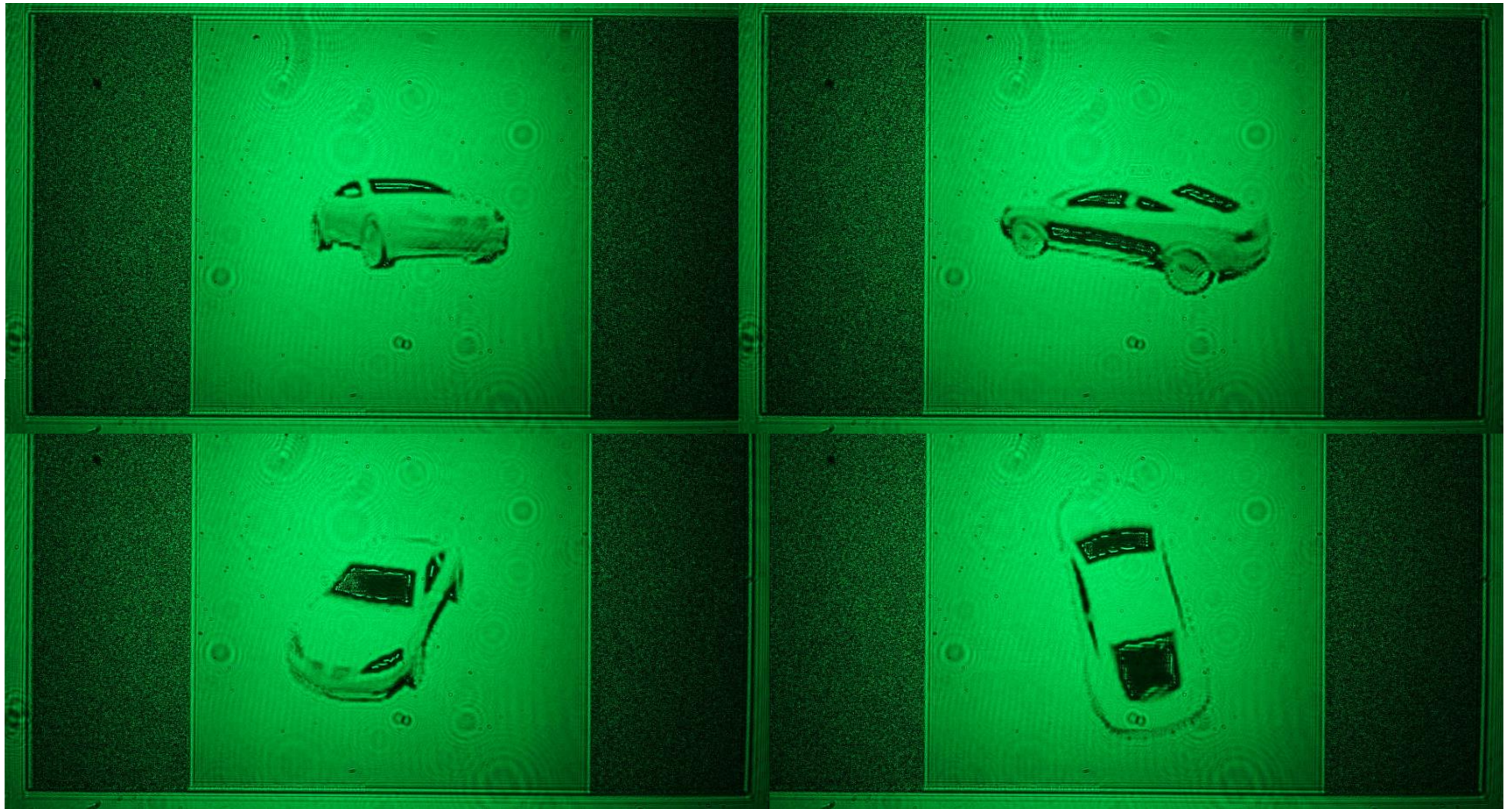
- Conventional NeRF requires multiple images (and their camera poses) taken from different angles of the same scene,
- Vision NeRF can infer 3D information about the scene from only one image and generate images from any viewpoint
- To generate a new viewpoint, the camera's position and angle is fed into Vision NeRF



Optical reconstructions

Single image

Hologram reconstruction

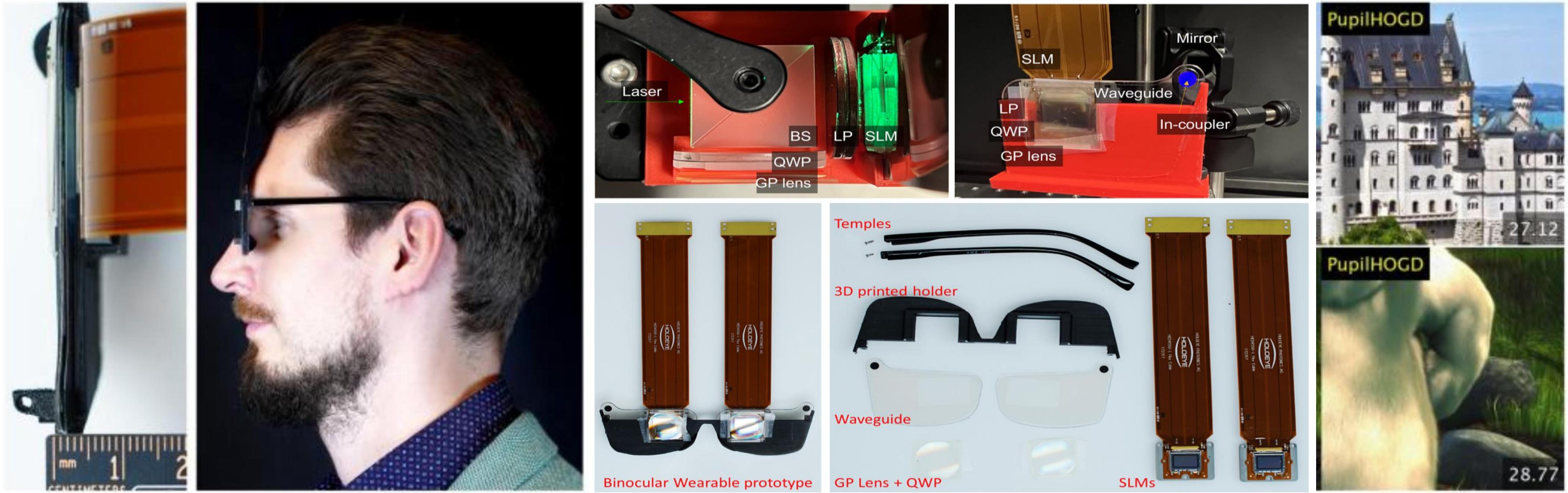


Topics

- Free viewpoint holographic displays using NeRF / 3DGS
- Free viewpoint holographic displays using Vision NeRF
- Binocular hologram generation
- 3D TV game using holographic signal converter

Holographic 3D glasses

Kim+ SIGGRAPH 2022



- ✓ Holographic 3D glasses offer true 3D hologram display, providing a more realistic and immersive augmented reality experience than traditional AR
- ✓ However, we need two 3D data and holograms for left and right eyes
- ✓ Time consuming

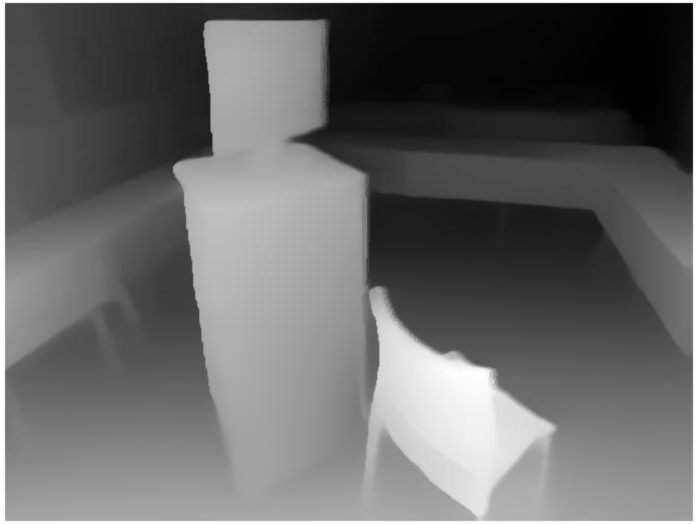
Binocular hologram generation using deep learning

Original image (for left eye)

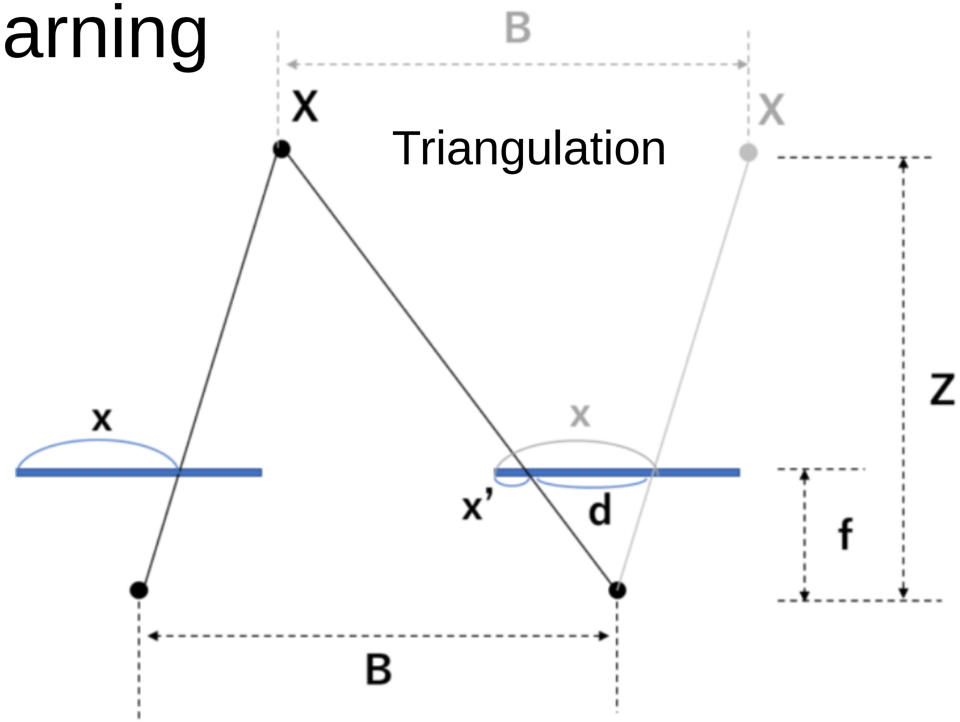
Binocular hologram generation using deep learning

Original image (for left eye)

Depth map



Triangulation

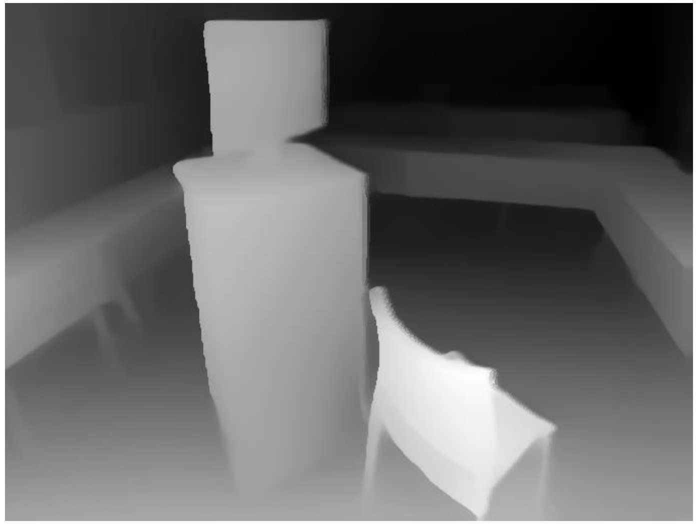


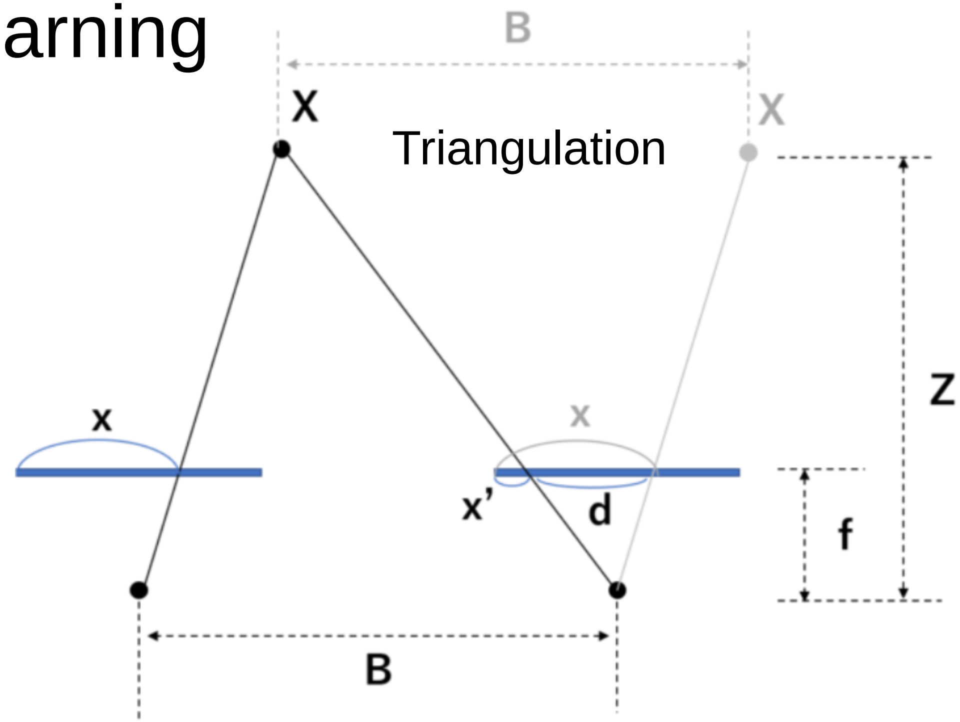
To generate a new image for the right eye, we use triangulation and a depth map to shift the pixels of the left eye's image to new positions of the right eye’s image

Binocular hologram generation using deep learning

Original image (for left eye)

Depth map



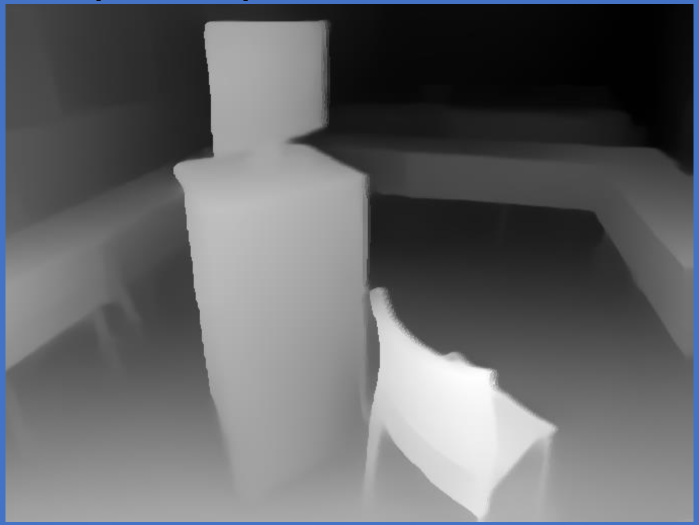


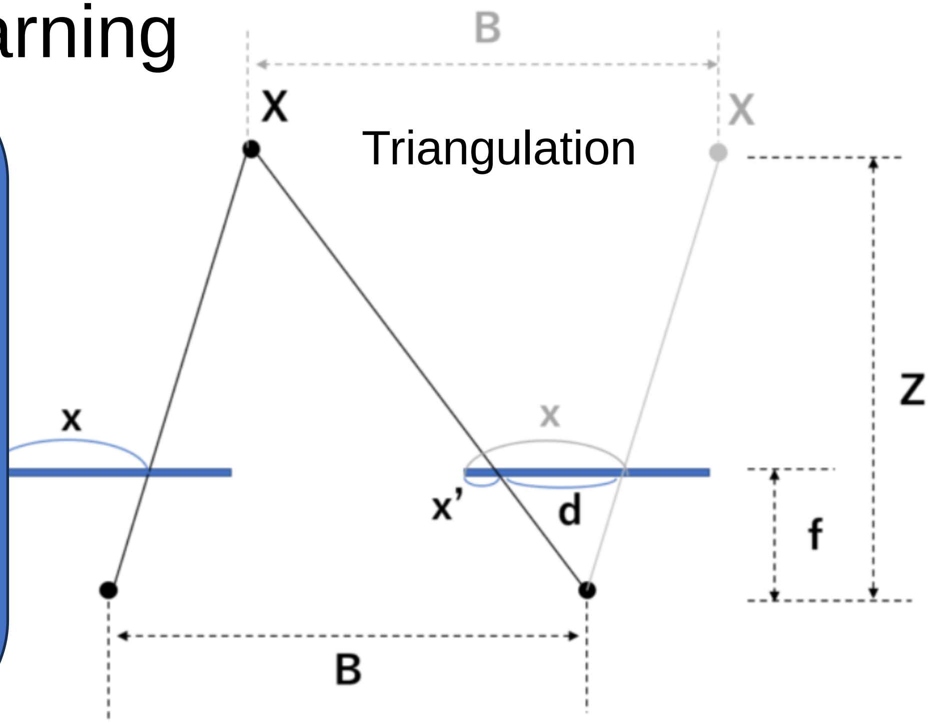
Deep neural network (gmcnn) for inpainting

Binocular hologram generation using deep learning

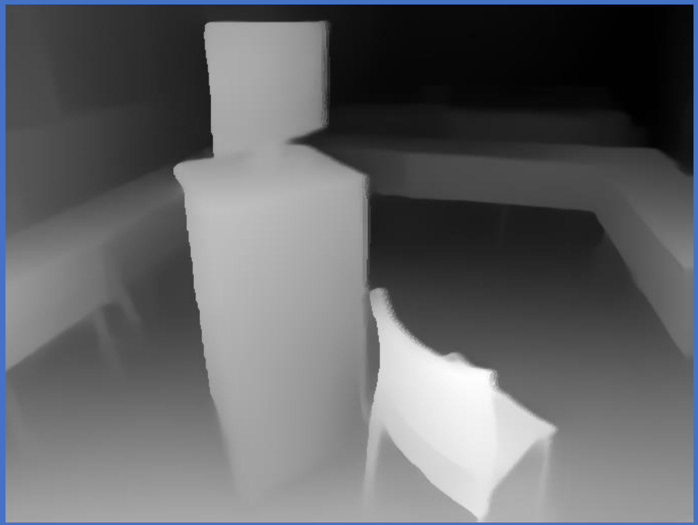
Original image (for left eye)

Depth map





Deep neural network (gmcnn) for inpainting



Hologram reconstruction for left eye

for right eye

Topics

- Free viewpoint holographic displays using NeRF / 3DGS
- Free viewpoint holographic displays using Vision NeRF
- Binocular hologram generation
- 3D TV game using holographic signal converter

Various video equipment

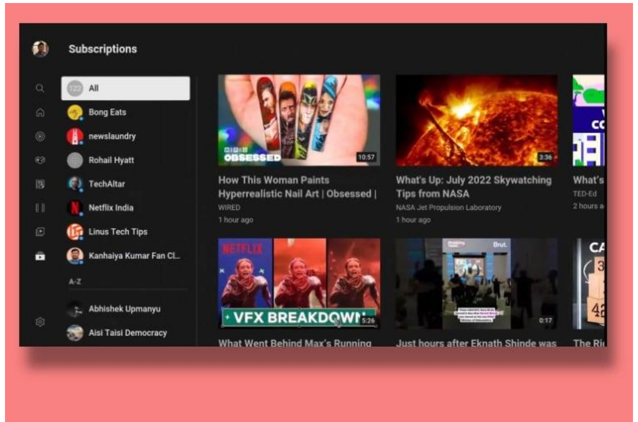
- We have many video equipment
  - ✓ TV, TV games, PC, YouTube, Amazon prime video…
  - ✓ Most video equipment presents two-dimensional movies

- If existing video equipment are modified to be able to display holographic 3D video, extensive equipment modifications would be required on both the transmitter and receiver sides

Nintendo Switch

Amazon prime video

Youtube



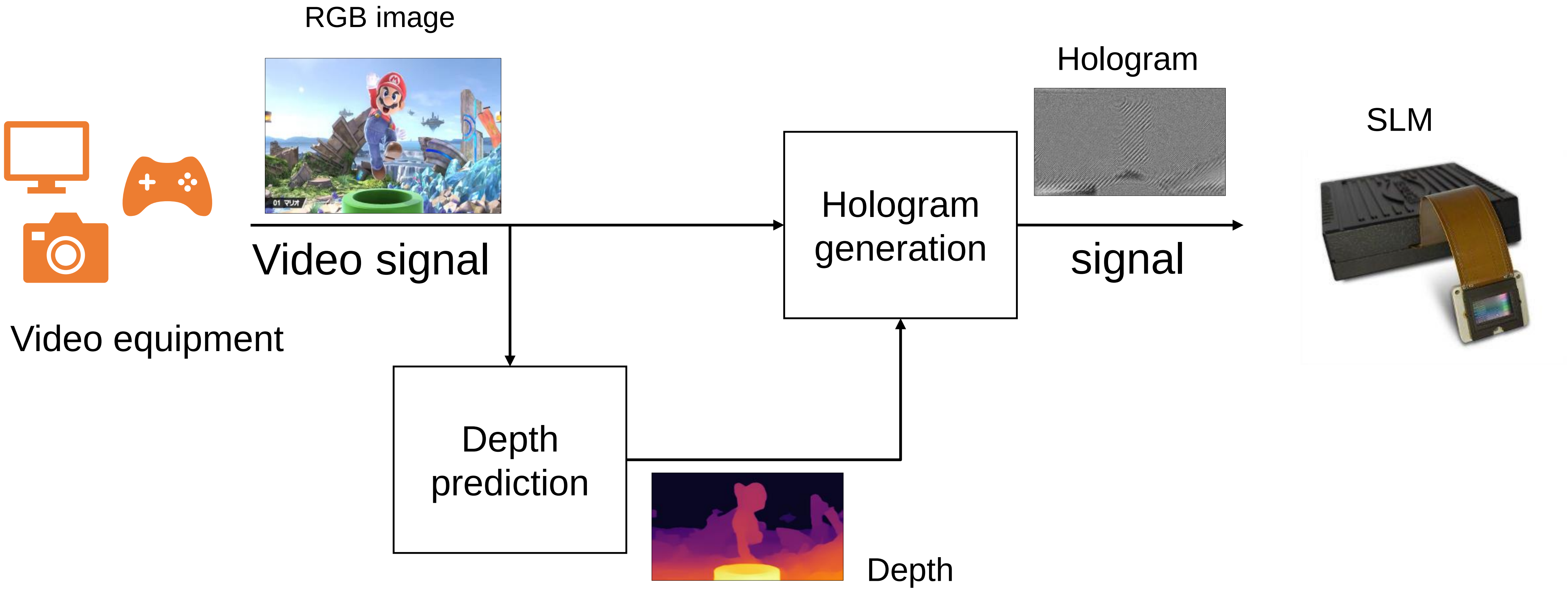
TV

Holographic signal converter

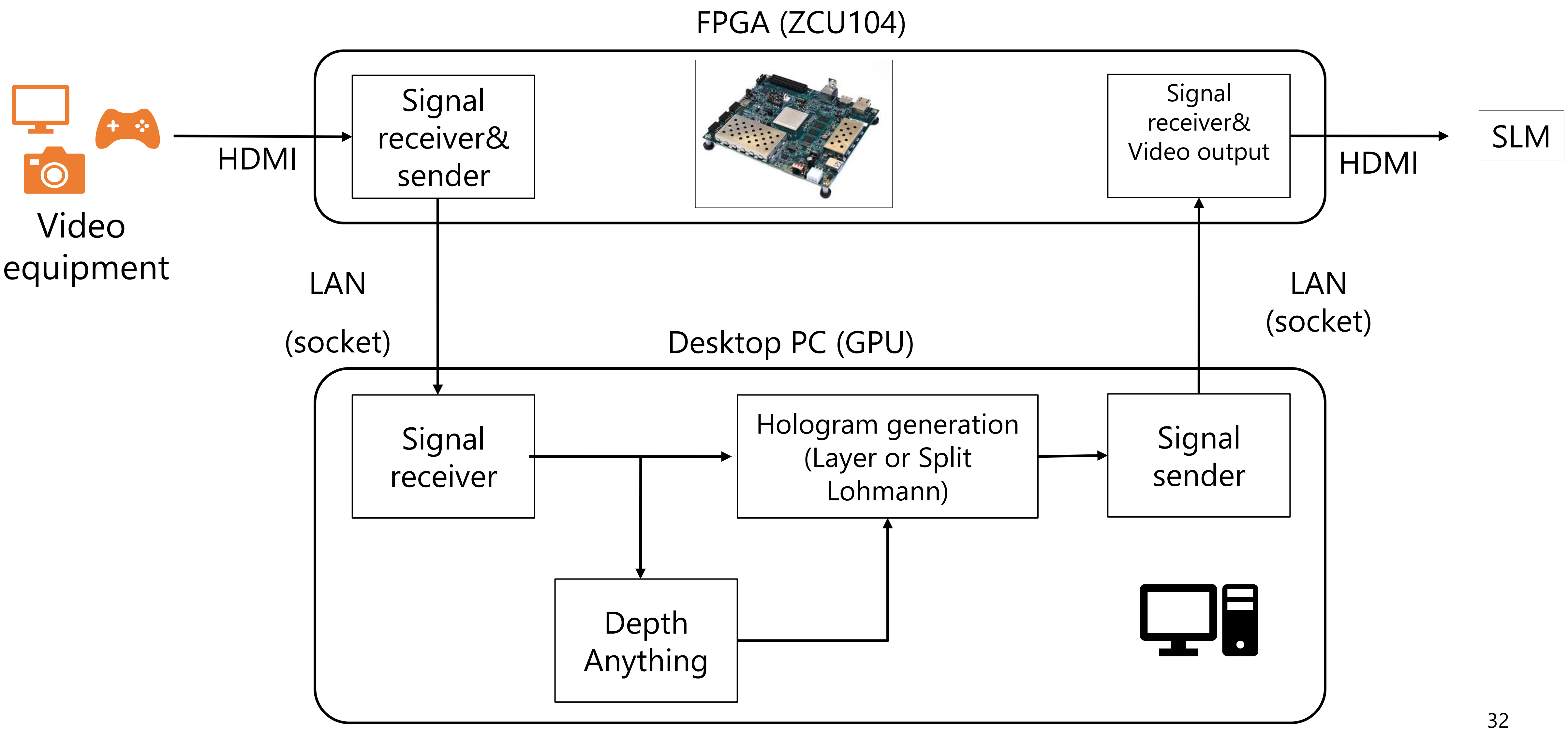
- We propose a conversion method from 2D signals of existing video equipment to holographic signal
- The proposed method can reproduce 3D holographic images in real-time
- The proposed method does not require any changes to existing video equipment

Masato Shotoku
Ph.D student

Proposed system: Holographic signal converter

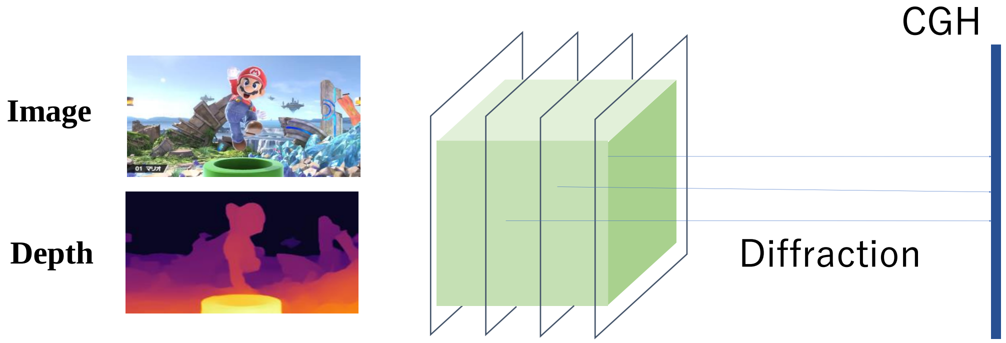


Proposed system

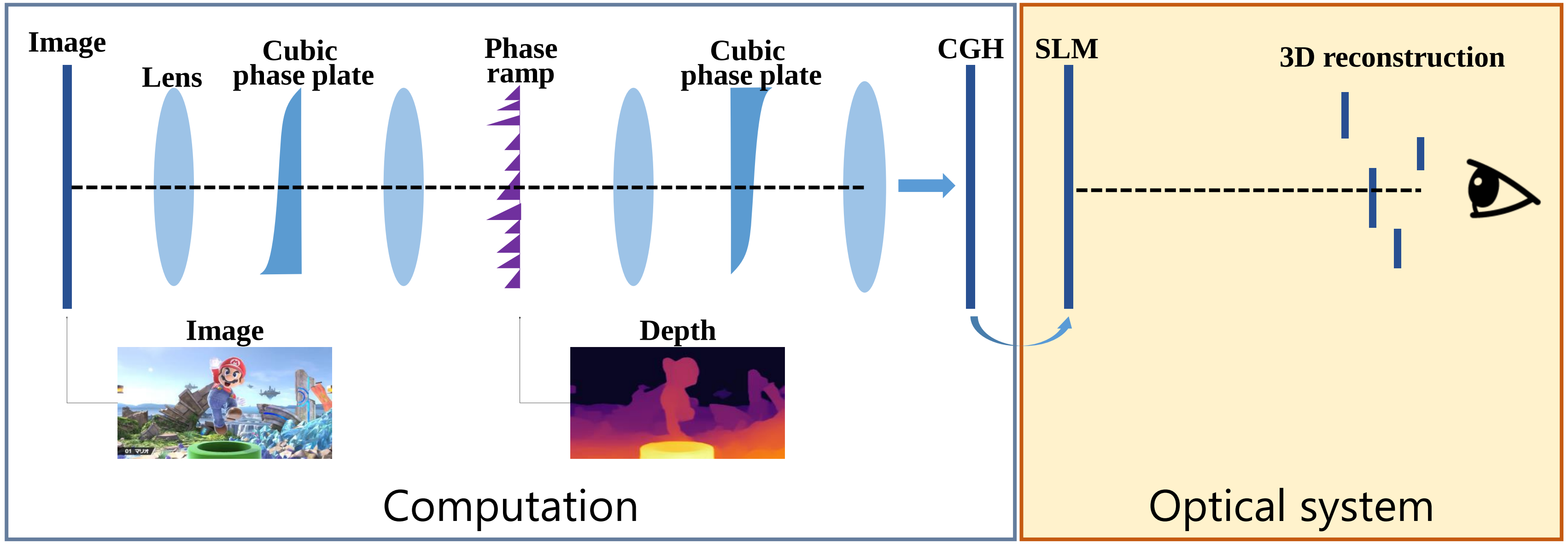


Hologram calculation algorithm

Layer method



Split Lohmann method

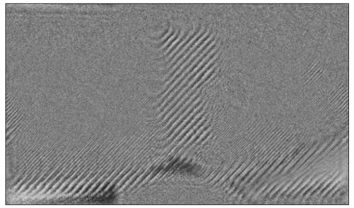


Chenliang+ Adv. Photon. Nexus 3 (2024)

Holographic TV game

Playing Mario Kart / GRAN TURISMO 7 on holographic display

Hologram

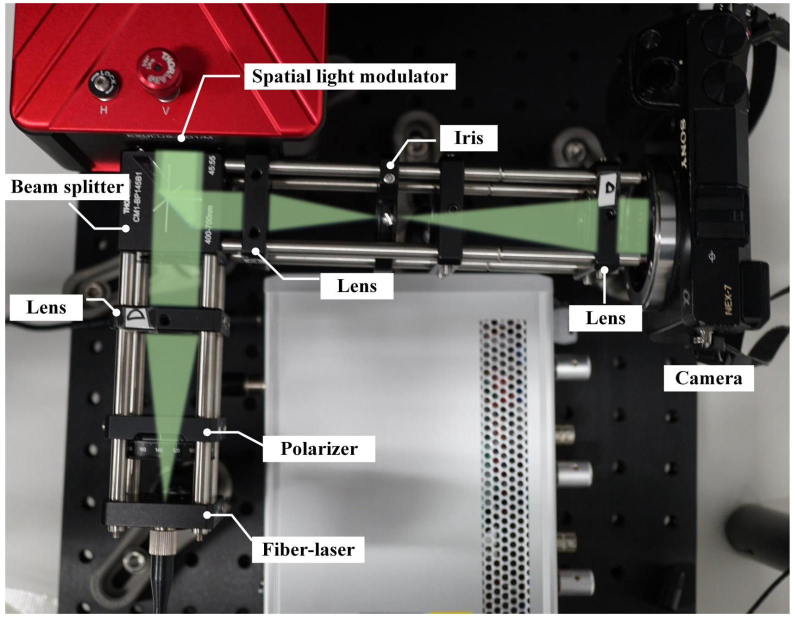


HDMI signal

Hologram calculation

Depth Estimation

Depth map



World's first holographic TV game?


06
3/3
1st

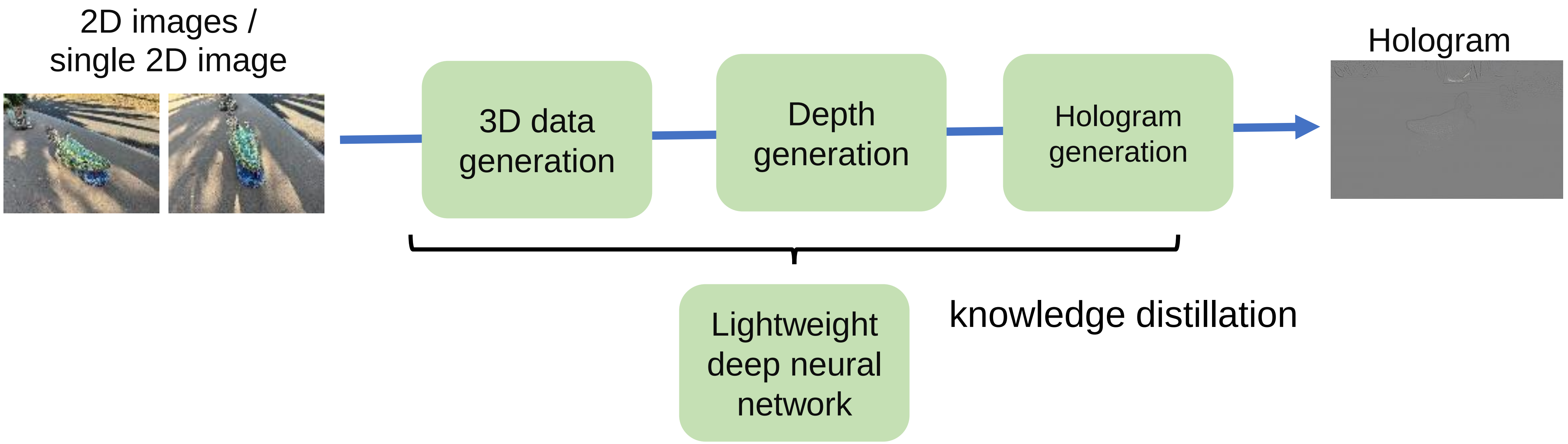
Shotoku, M., Wang, F., Ito, T., & Shimobaba, T. Holographic signal converter for existing two-dimensional video equipment (under review)

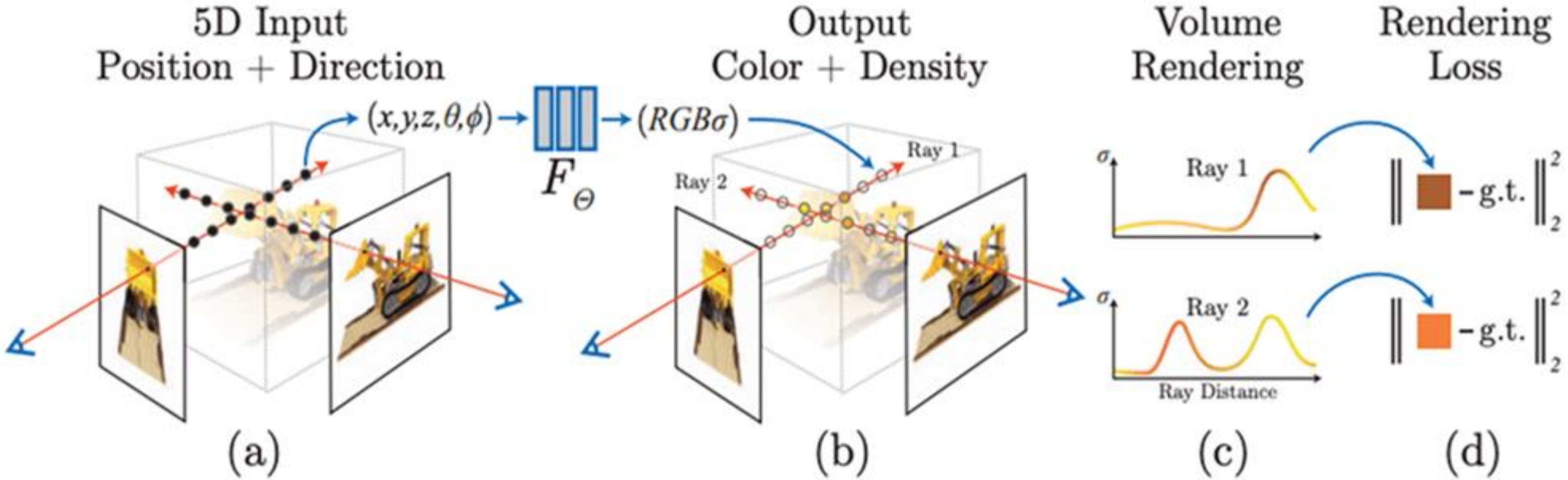
Sony PlayStation 5 GRAN TURISMO 7

We would like to express our gratitude to Polyphony Digital for the permission to use GRAN TURISMO 7.

Conclusions

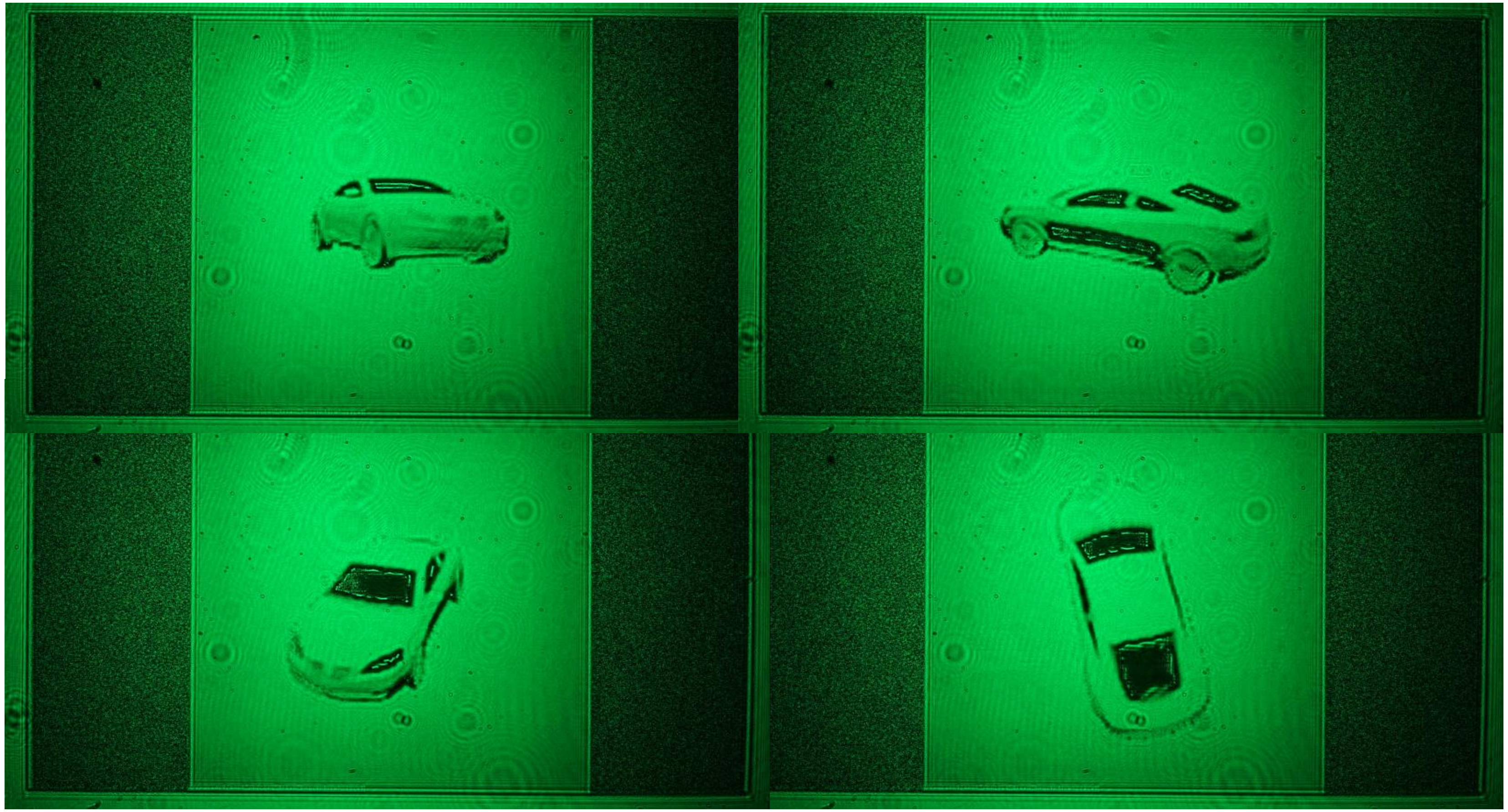
- We introduced the following topics:
  - Free viewpoint holographic displays using NeRF
  - Free viewpoint holographic displays using Vision NeRF
  - Binocular hologram generation
  - 3D TV game using holographic signal converter
- Future works
  - Knowledge distillation + hologram generator for deep 3D scene




5D Input
Position + Direction
(x,y,z,θ,ϕ)
F_Θ
(RGBσ)
Output
Color + Density
Ray 1
Ray 2
Volume
Rendering
Rendering
Loss
σ
Ray 1
σ
Ray 2
Ray Distance
- g.t.
- g.t.
(a)
(b)
(c)
(d)

Optical reconstructions

ホログラムの仕様			
解像度	1920×1080[pixel]	光波長	G:520[nm]
画素ピッチ	6.4[μm]	伝搬距離	16[cm]



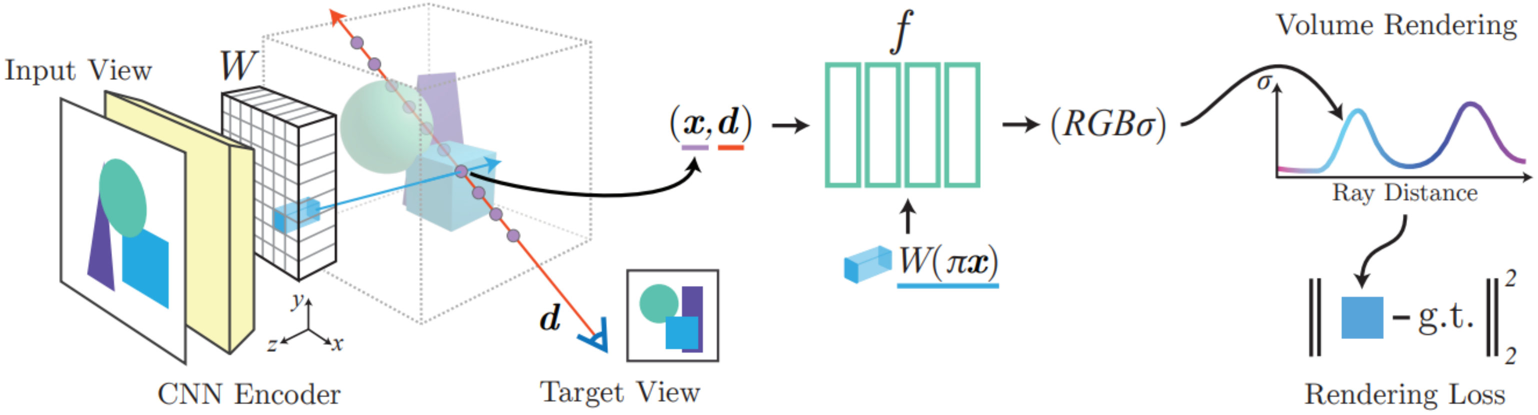
提案手法

STEP 02

元

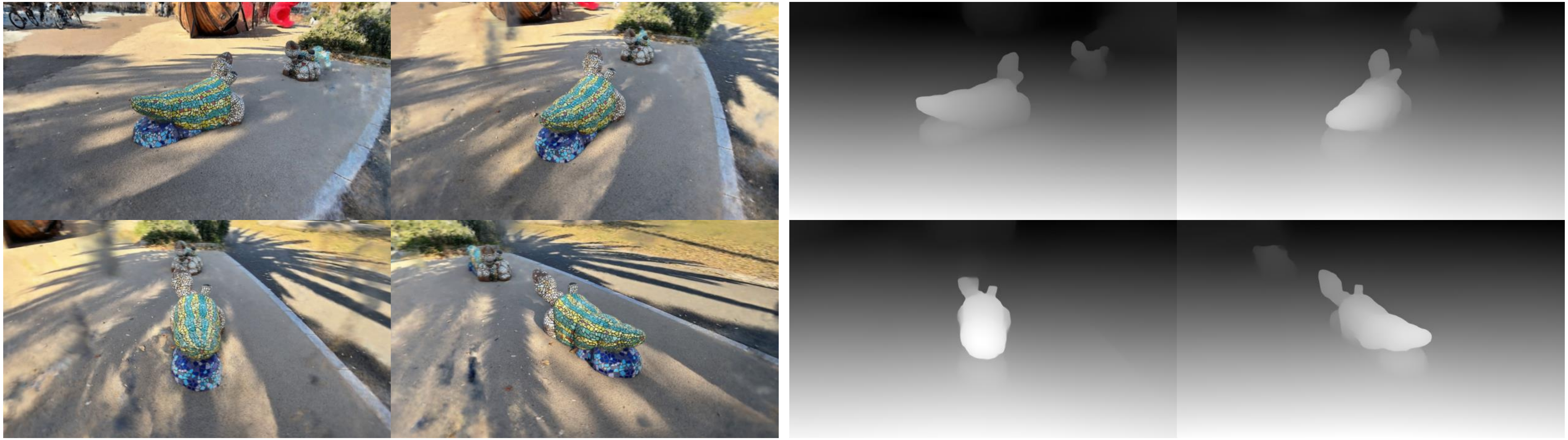
Pixel NeRF：NeRFを応用して入力画像が単数や少数の場合でも
３次元シーンの再構成を可能にした技術

学習時にCNNで特徴量を抽出し，どのシーンにも
共通する情報として使用



図の引用：Alex Yu, Vickie Ye, Matthew Tancik, Angjoo Kanazawa.
pixelNeRF: Neural Radiance Fields from One or Few Images
arxiv:2012.02190

Result of depth prediction



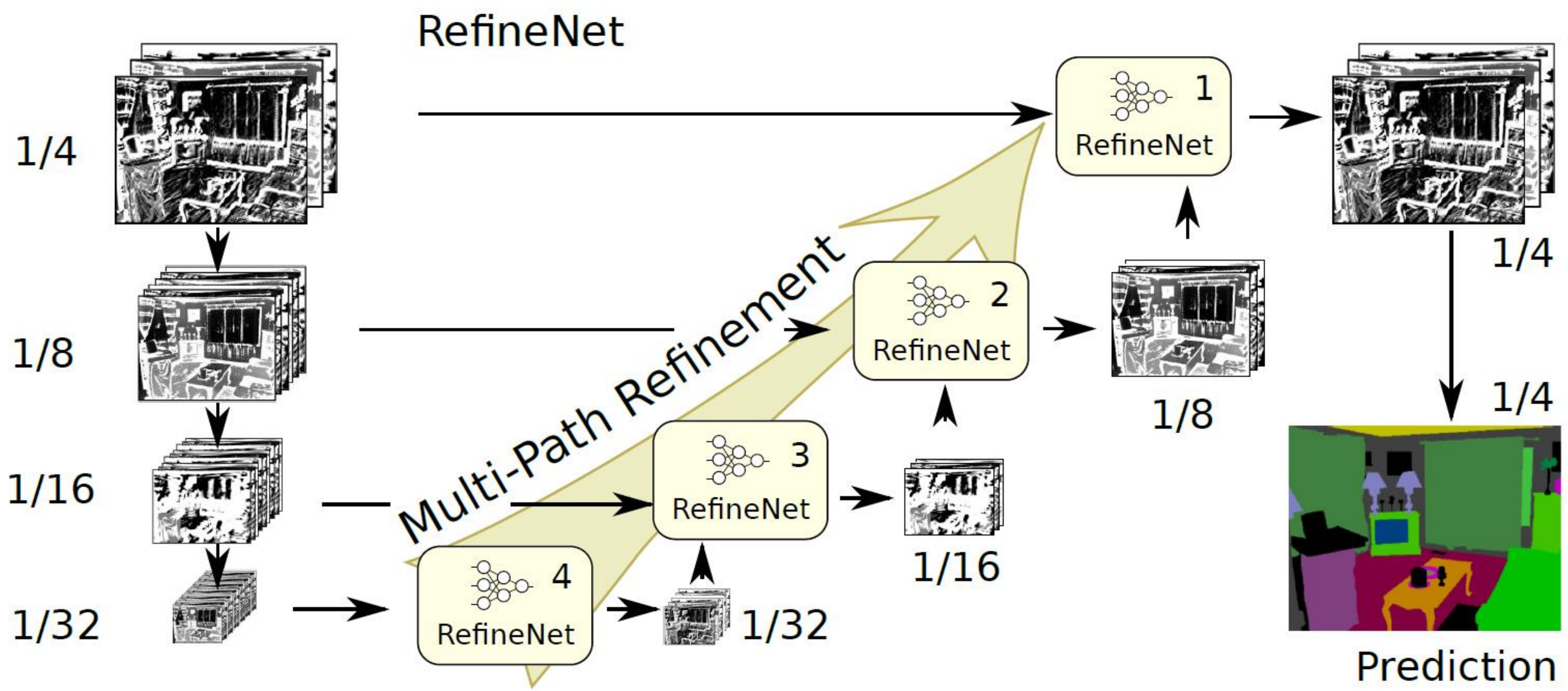
Depth prediction: MiDaS

Network structure: RefineNet

End-to-end learning

Over 10 diverse datasets used

$$d = \mathcal{N}_{\text{MiDaS}}\{\text{RGB}\}$$



3D scene reconstruction from NeRF

$$RGB = \mathcal{N}_{NeRF}\{\mathbf{v}\}$$

Elapsed times in each step

	Split Lohmann		Layer-based		
	Gray	Color	Gray	Color	
HDMI extraction	1.10	1.46	1.09	2.23	ms
Send to GPU	4.48	10.2	4.41	10.8	
Depth Anything	20.9	22.1	21.6	21.8	
Hologram generation	2.92	9.24	10.1	28.7	
Send to GPU	15.5	**96.2**	16.0	**96.0**	
HDMI conversion	0.83	0.94	0.79	0.91	
FPS	21.6	6.32	20.1	5.62	fps
FPS (Pipeline)	51.3	9.49	46.5	9.52	

Elapsed times in each step

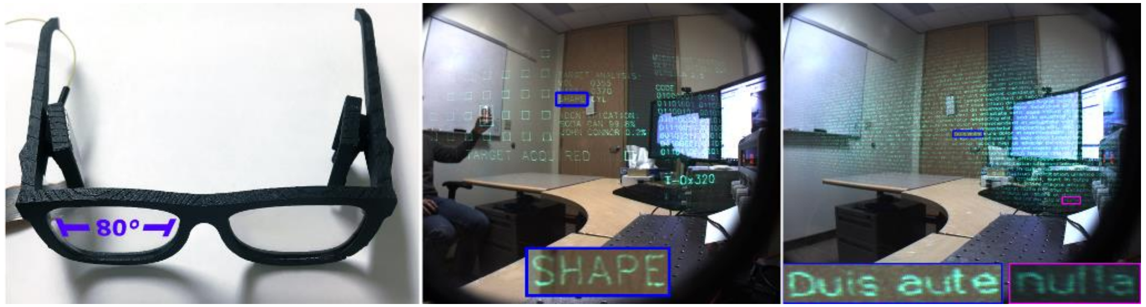
My expectation:

We need a fast and efficient compression/decompression technique

Could JPEG Pleno distribute reference codes for the compression algorithms?

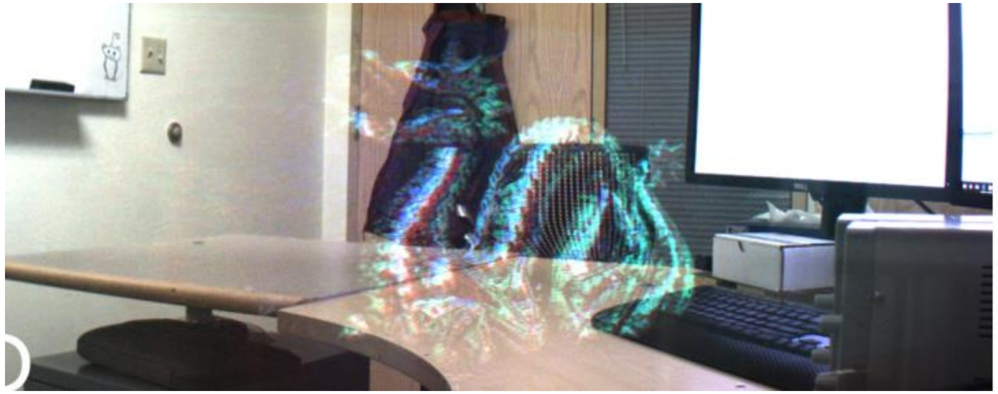
	Split Lohmann				
	Gray	Co			
HDMI extraction	1.10	1.4			
Send to GPU	4.48	10.2	4.41	10.8	ms
Depth Anything	20.9	22.1	21.6	21.8	
Hologram generation	2.92	9.24	10.1	28.7	
Send to GPU	15.5	**96.2**	16.0	**96.0**	
HDMI converter	0.83	0.94	0.79	0.91	
FPS	21.6	6.32	20.1	5.62	fps
FPS (Pipeline)	51.3	9.49	46.5	9.52	

Near-eye holographic display



https://www.microsoft.com/en-us/research/project/holographic-near-eye-displays-virtual-augmented-reality/

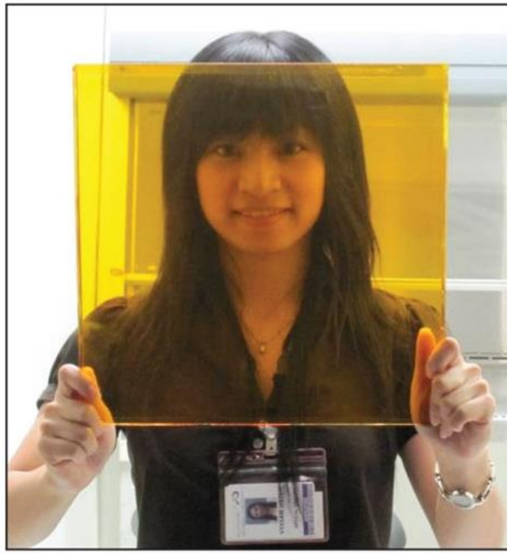
Practical use within 10 years?



Maimone+ Trans Graph 2017

Holographic display with wide angle and large image size

e.g. Photorefractive polymer (rewritable holographic polymer)

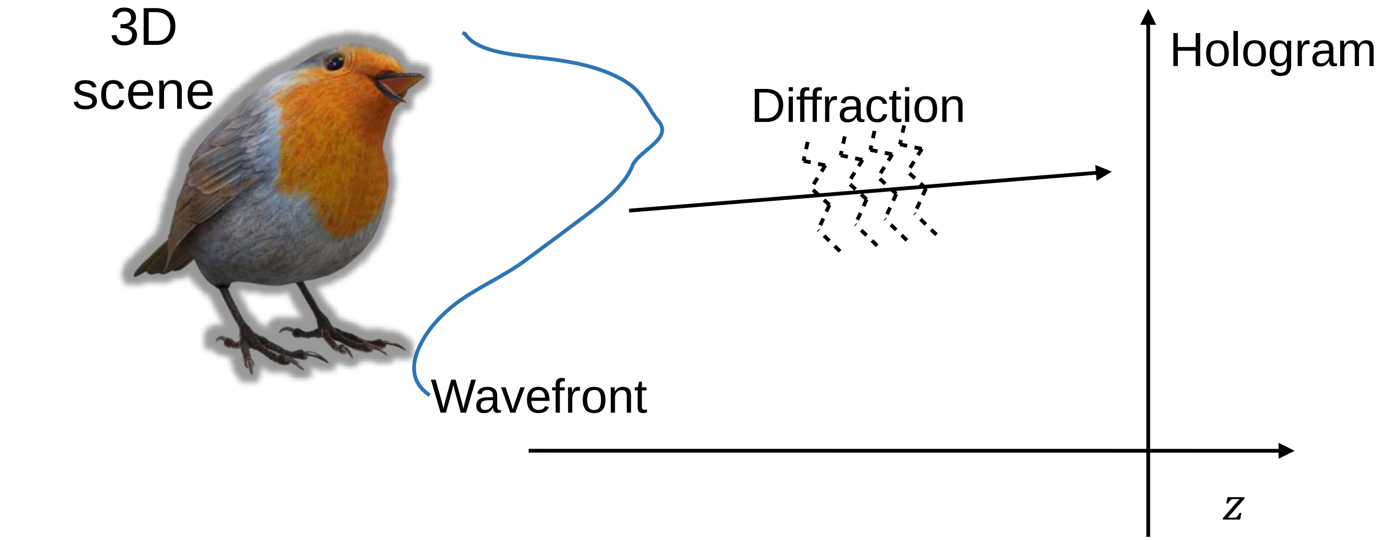


Large space bandwidth product

It would still take over 30 years to make it practical?

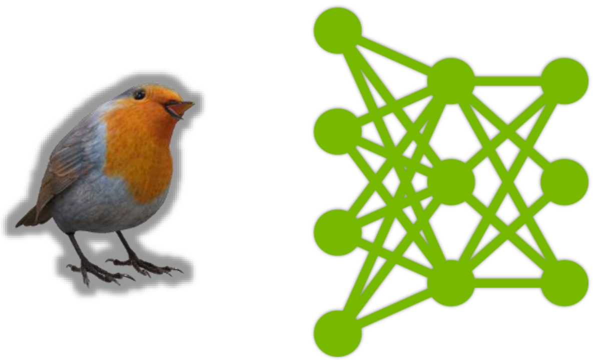
Blanche+ Nature 2010

Acceleration algorithms

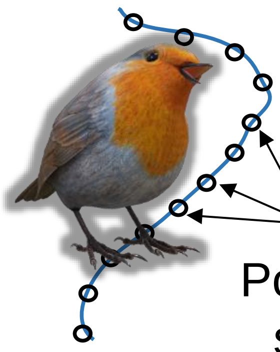


Deep learning

Deep neural network can directly predict holograms from 3D scene

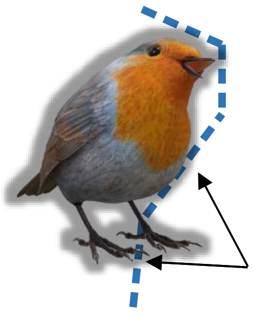


Point cloud



3D scene is expressed as the aggregation of point light sources

Polygon



3D scene is expressed as the aggregation of polygons

Light field method

Multi-view image

Layer (RGB +Depth image)

Point cloud method

Suitable for wide-angle holographic display

$O(N^4)$

- LUT
  - Novel LUT, Kim+ Appl.Opt. (2008)
  - Mini LUT using principal component analysis (PCA), Jiao+ Opt. Express (2017)
- Recurrence formulas
  - Yoshikawa+ Opt. Rev 2001, Matsushima+ Appl. Opt. 2000, Shimobaba+ Comput. Phys. Commun. 2001
- Separable convolution
  - S-LUT method, Pan+ Opt. Express (2009)
  - Separable convolution, Maimone+ ACM Trans. Graph. (2017)
  - Oriented separable convolution, Shimobaba+ Opt. Express (2022)

$O(N^3)$

- Sparsity
  - Image hologram, Yamaguchi+ Opt. Eng. (2007)
  - WRP, Shimbaba+ Opt. Lett. (2009)
  - Wavelet, Shimobaba+ Opt. Express (2017)
  - Short-time Fourier transform
    - Yamaguchi+, Yoshikawa+, Blinder+
  - Gabor transform, Blinder+ Photo. Res. (2025) $O(N)$!!

Polygon method

- Rotational diffraction (FFT-based method)
  - Matsushima+ J. Opt. Soc. Am. 2003
  - Matsushima+ Opt. Express 2014
  - Nishi and Matsushima Appl. Opt. 2017

Wide-angle holographic display

- Analytical solution method
  - Ahrenberg+ Appl. Opt. 2008
  - Liu+ Opt. Express 2010
  - Liu+ Opt. Lett. 2011
  - Pan+ Appl. Opt. 2013
  - Zhang+ Opt. Express 2018
  - Wang+ Photo. Res. 2023
  - Wang+ Opt. Lasers Eng. 2023

Near-eye holographic display

- Hybrid approach combining polygon with point cloud
  - Wang+ Opt. Lett. 2023

Wide-angle holographic display

Prof. Fan Wang
Chiba Univ.

Light field method

- Holographic stereogram
  - Yatagai+ Appl. Opt. 1976
- Phase-added stereogram (PAS)
  - Yamaguchi+ SPIE Conf. 1993
- Ray sampling plane method
  - Wakunami+ Opt. Express 2011

Wide-angle holographic display

- Layer method
  - Multiple diffraction calculations
    - Shimobaba+ Opt. Express 2013, Chen+ Opt. Express 2015
  - Split-Lohmann holography
    - Chang+ Adv. Photo. Nexus 2024

Near-eye holographic display

  - Hybrid approach combining layer method + Stereogram
    - Zhang+ Opt. Express 2015

Wide-angle holographic display

Light field method

- Holographic stereogram
  - Yatagai+ Appl. Opt. 1976
- Phase-added stereogram (PAS)
  - Yamaguchi+ SPIE Conf. 1993
- Ray sampling plane method
  - Wakunami+ Opt. Express 2011

Wide-angle holographic display

- Layer method
  - Multiple diffraction calculations
    - Shimobaba+ Opt. Express 2013, Chen+ Opt. Express
  - Split-Lohmann holography
    - Chang+ Adv. Photo. Nexus 2024
  - Hybrid approach combining layer method + Ster
    - Zhang+ Opt. Express 2015

My expectation:

Many CGH acceleration algorithms have been proposed since the 1970s

Could JPEG Pleno collect reference codes for the various algorithms?

CGH algorithm’s zoo

Reference codes need not be optimized

Hardware accelerations

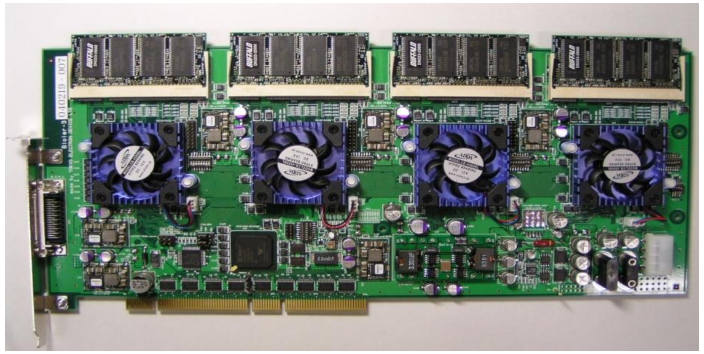
Multi-core CPU

Parallel processors

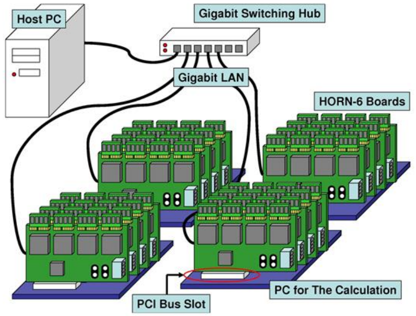
GPU

XeonPhi coprocessor

FPGA (Field Programmable Gate Array)



HORN-5 board (2004)



HORN-6 Cluster (2009)

HORN-8 (2018)

HORN : HOlographic ReconstructiON

Special-purpose computers for holography: HORN

Prof./Dean Tomoyoshi Ito,
Chiba Univ.
Originator of HORN project

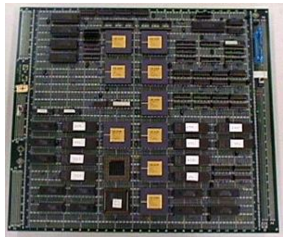
HORN-1

IC
1 circuit / 1board
T.Ito

1993

HORN-2

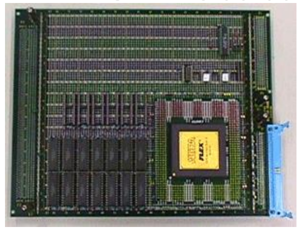
DSP
T. Ito



1994

HORN-3

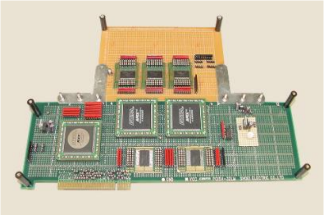
FPGA
1 circuit / 1 chip
T. Shimobaba



1998

HORN-4

FPGA
New algorithm
20 circuits/ 1chip, T. Shimoabba

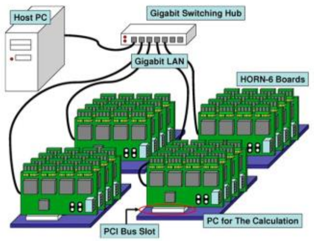


2001

HORN-5

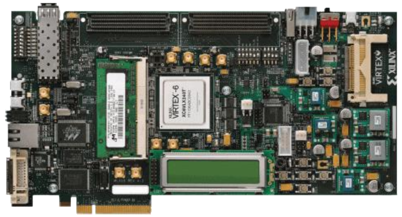
FPGA (PCB)
2005

HORN-6



FPGA (cluster)
Y. Ichihashi(NICT)
2009

HORN-7

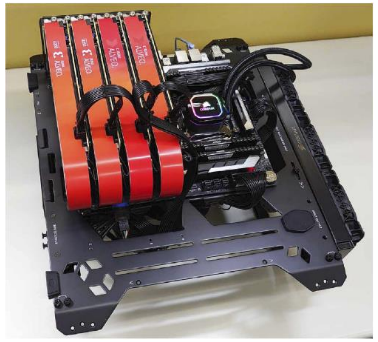


FPGA (Xilinx board)
N. Masuda, Y. Endo
Phase-only hologram
2012

HORN-8

FPGA (PCB)
2015

HORN-9

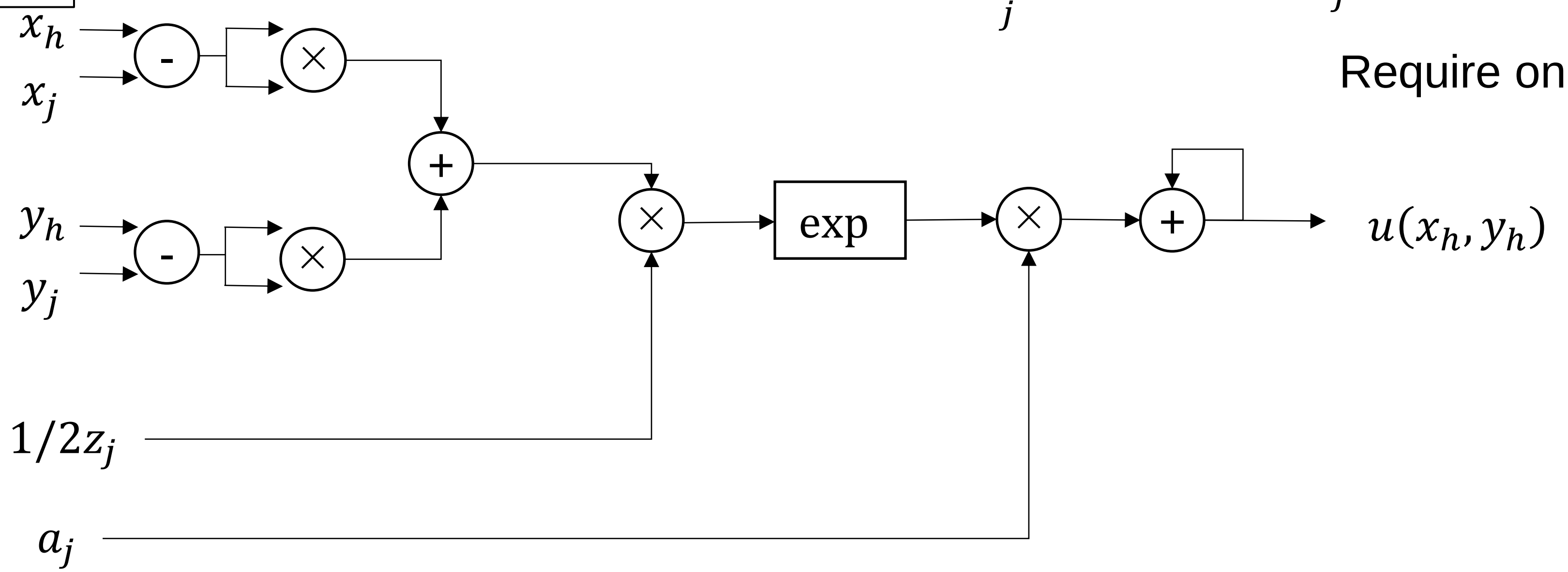


FPGA boards
Yamamoto
Hilbert transform
2022

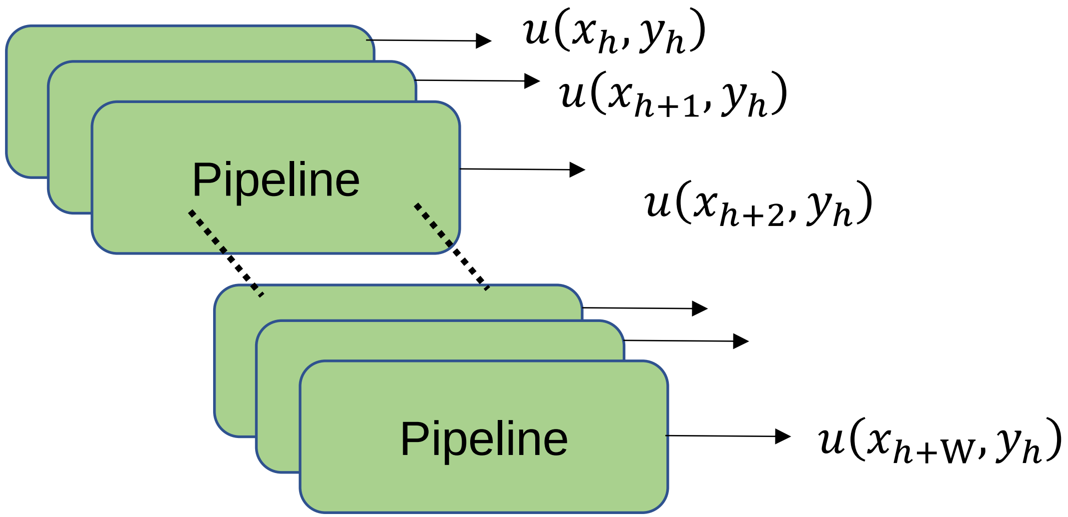
Pipeline architecture

$$u(x_h, y_h) = \sum_{j}^{N} a_j \exp(ik\frac{x_{hj}^2 + y_{hj}^2}{2z_j})$$

Require only N clock cycle



Parallelizing pipelines

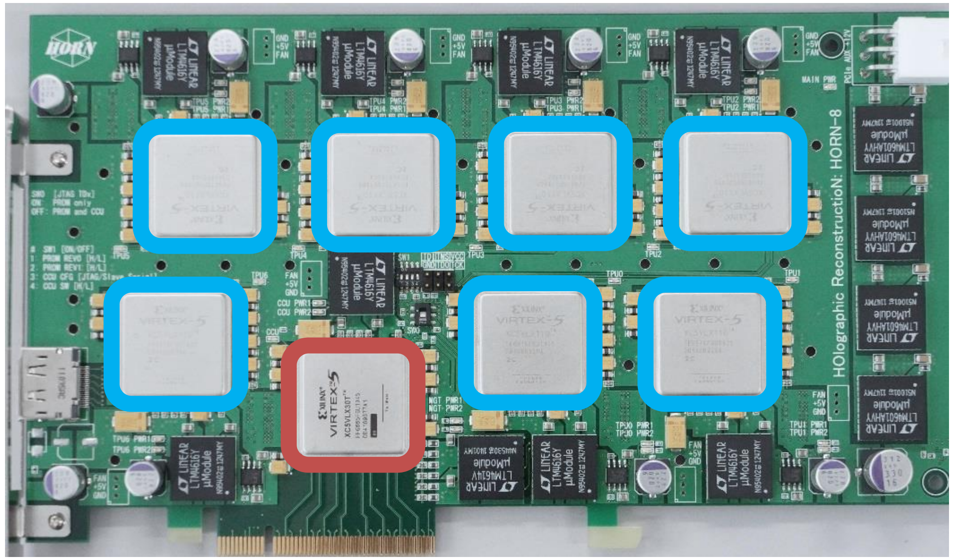


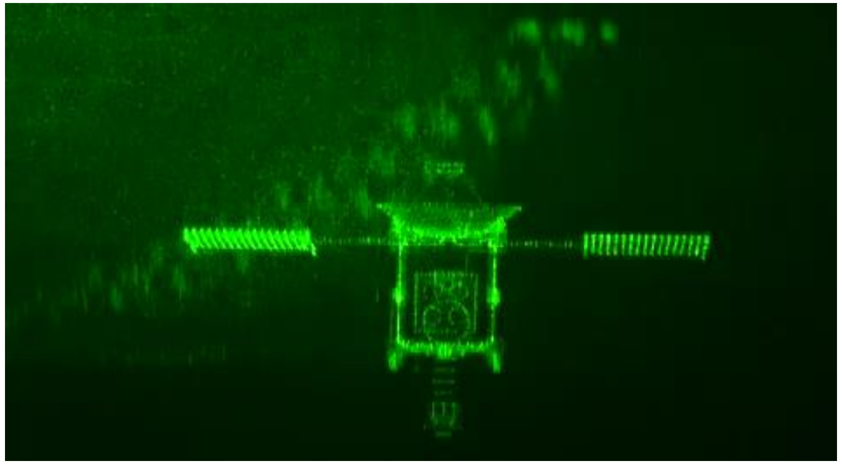
Multiple hologram pixels can be calculated in parallel

HORN-8

Nature Electronics, 1, 254–259 (2018)

$$u(x_h, y_h) = \sum_{j}^{N} a_j \exp(ik\frac{x_{hj}^2 + y_{hj}^2}{2z_j})$$





Hologram resolution is 1920x1080 (2K)

- Large-scale FPGA board developed in our laboratory
- 1 Control FPGA and 7 computational FPGAs
- Implementing about 4500 pipeline circuits in the FPGA board
  - 4500 hologram pixels can be calculated in parallel

Table 1 | Performance comparison of the HORN-8 board, CPU and GPU

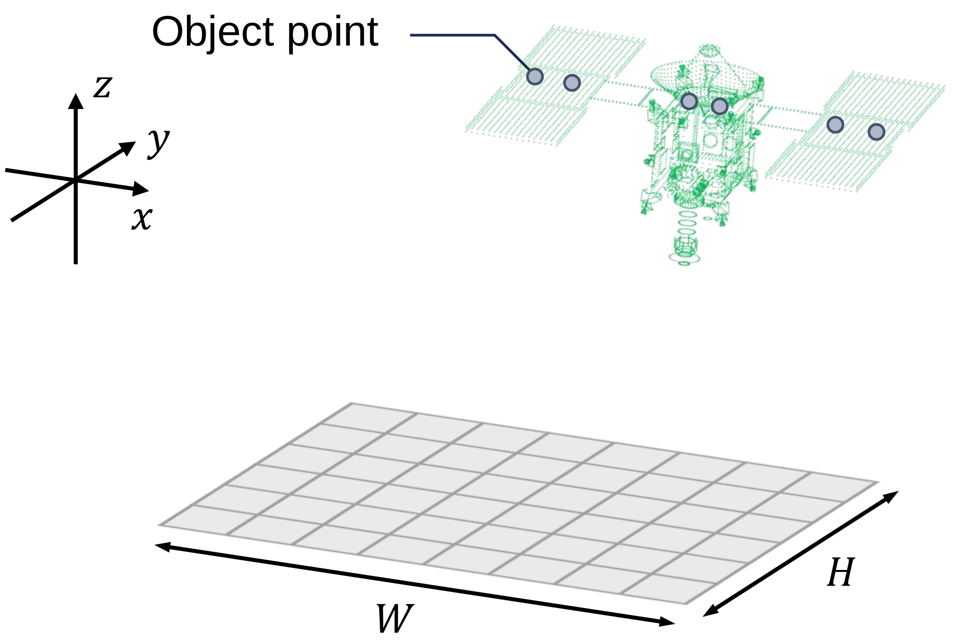
System	Time/CGH (s)	Speed ratio	Frame rate (fps)
CPU: Core i7-6700K, 4 cores, 4 GHz	2.951	1	0.34
GPU: GTX TITAN X, 3,072 cores, 1 GHz	0.106	28	9.4
HORN-8 board: 4,480 parallel (cores), 0.25 GHz	0.019	155	53

This is the time taken to generate a hologram of 1,920 × 1,080 pixels from a 3D image of 10,000 points.

HORN-10 (under development)

Prior to HORN-10

$$U(x_h, y_h) = \sum_{j=1}^{N_X N_Y} \exp\left\{ i \frac{\pi}{\lambda z_l} \left(x_{hj}^2 + y_{hj}^2 \right) \right\}$$

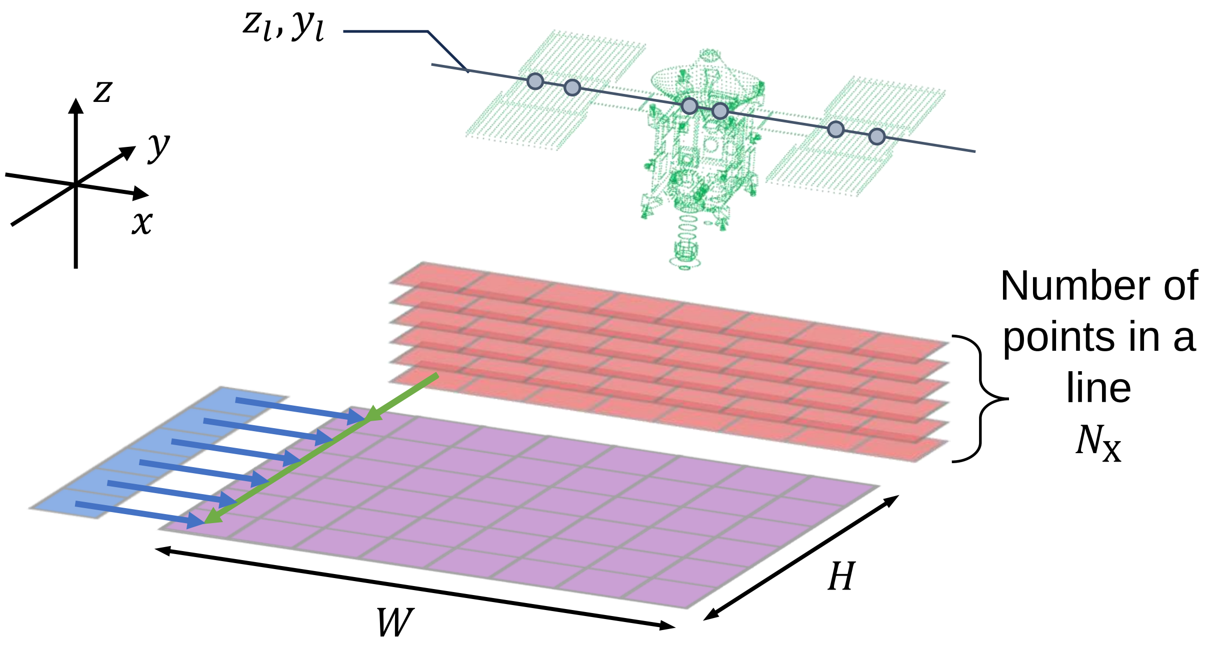


Computational complexity

$$O(N_X \times N_Y \times H \times W) = O(n^4)$$

HORN-10 (Separable convolution)

$$U(x_h, y_h) = \sum_{l=1}^{N_Y} \left\{ \exp\left(i \frac{\pi}{\lambda z_l} y_{hl}^2 \right) \sum_{j=1}^{N_X} \exp\left(i \frac{\pi}{\lambda z_l} x_{hl_j}^2 \right) \right\}$$



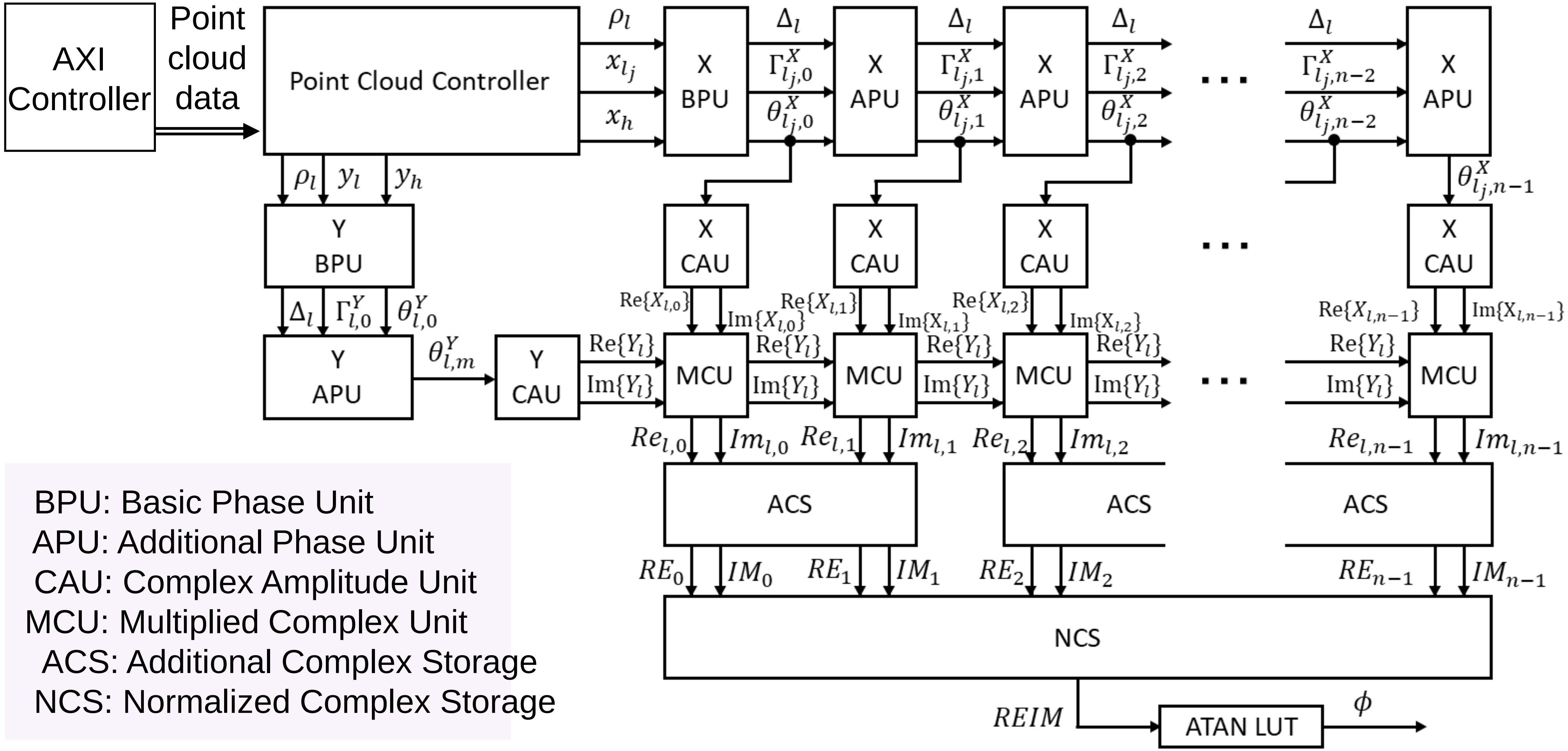
Computational complexity

$$O(((N_X + H) \times W + H) \times N_Y) = O(n^3)$$

Block diagram of HORN-10

$$U = \sum_{l=1}^{N_Y} \left\{ \exp\left(i \cdot 2\pi\theta_l^Y\right) \sum_{j=1}^{N_{X,l}} \exp\left(i \cdot 2\pi\theta_{l_j}^X\right) \right\}$$

■ 1920 pipelines



BPU: Basic Phase Unit
APU: Additional Phase Unit
CAU: Complex Amplitude Unit
MCU: Multiplied Complex Unit
ACS: Additional Complex Storage
NCS: Normalized Complex Storage

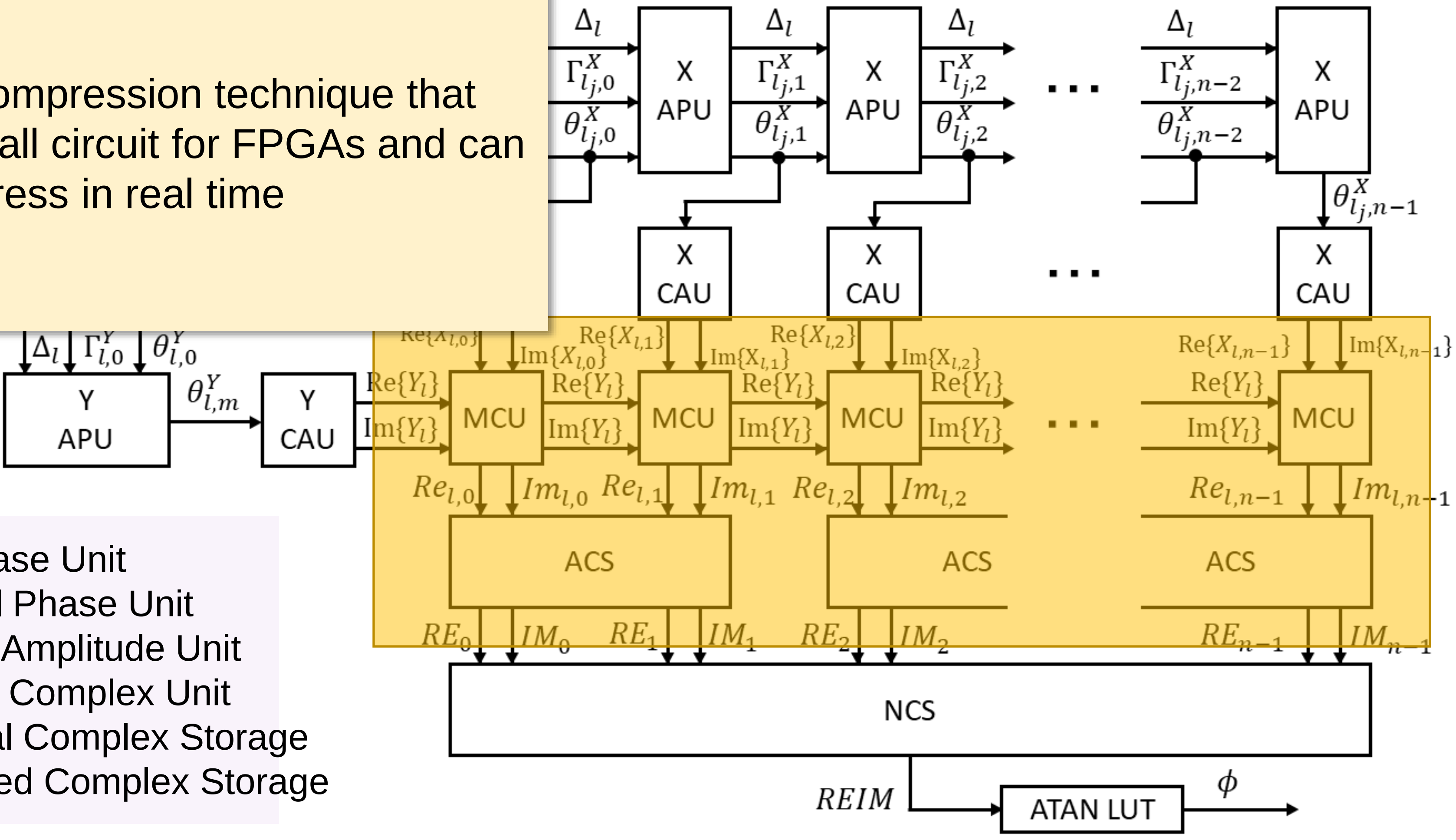
Block diagram of HORN-10

$$U = \sum_{l=1}^{N_Y} \left\{ \exp\left(i \cdot 2\pi\theta_l^Y\right) \sum_{j=1}^{N_{X,l}} \exp\left(i \cdot 2\pi\theta_{l_j}^X\right) \right\}$$

My expectation

Current pipeline count is limited by FPGA memory capacity

We need a hologram compression technique that can be realized in a small circuit for FPGAs and can compress and decompress in real time



BPU: Basic Phase Unit
APU: Additional Phase Unit
CAU: Complex Amplitude Unit
MCU: Multiplied Complex Unit
ACS: Additional Complex Storage
NCS: Normalized Complex Storage

Conclusion

- We presented our current studies
- We also described my expectations for JPEG Pleno for the problems in those studies:
  - Standard implementation for CGH algorithms
  - How to evaluate image quality without a reference image
  - Techniques for compressing and decompressing holograms at high speed


千葉大学


Tokyo
Narita Airport
Haneda Airport
Tokyo Disney Land
Chiba University
千葉大学 西千葉キャンパス