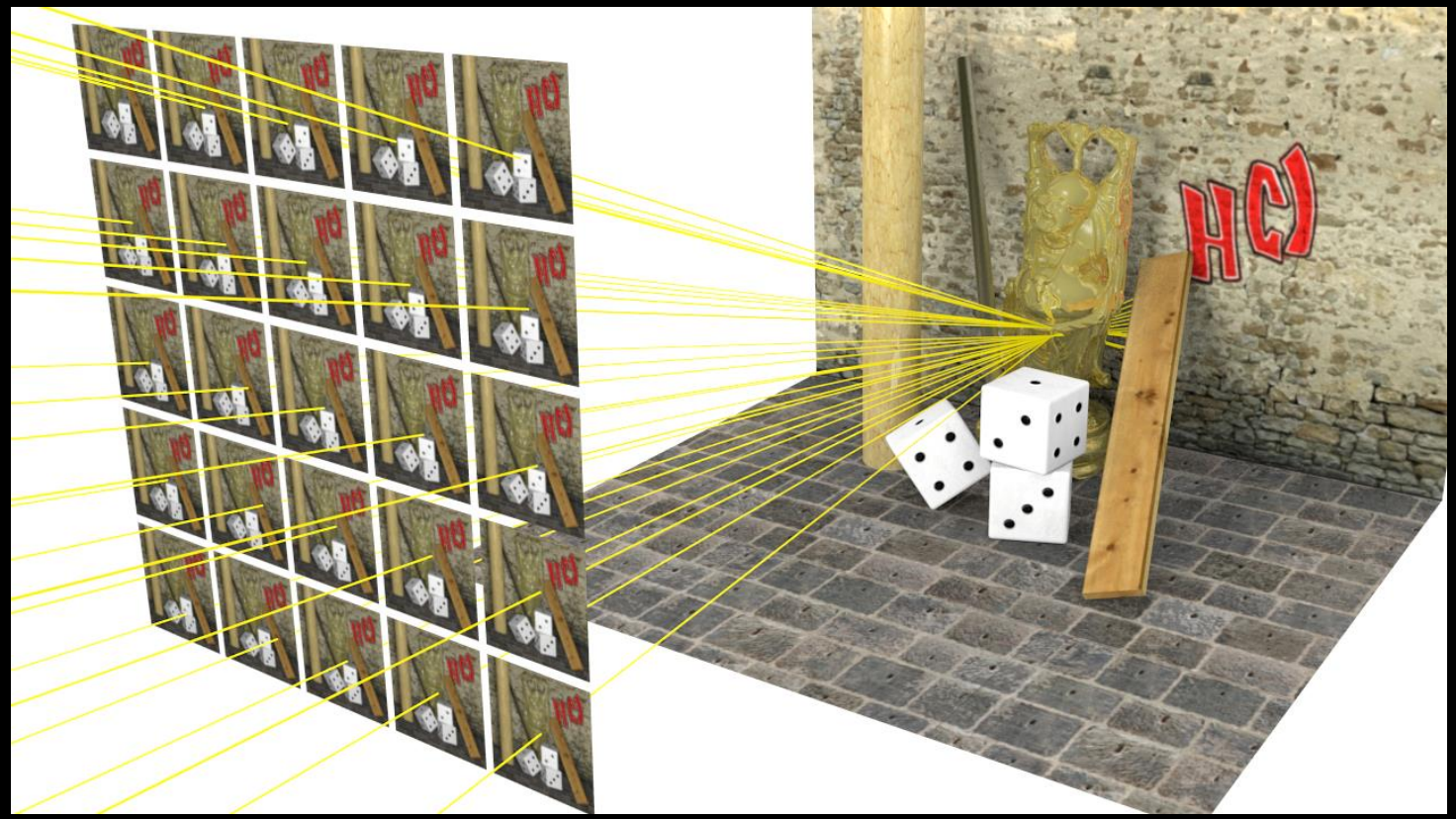
上海科技大学
ShanghaiTech University

Surface Light Fields

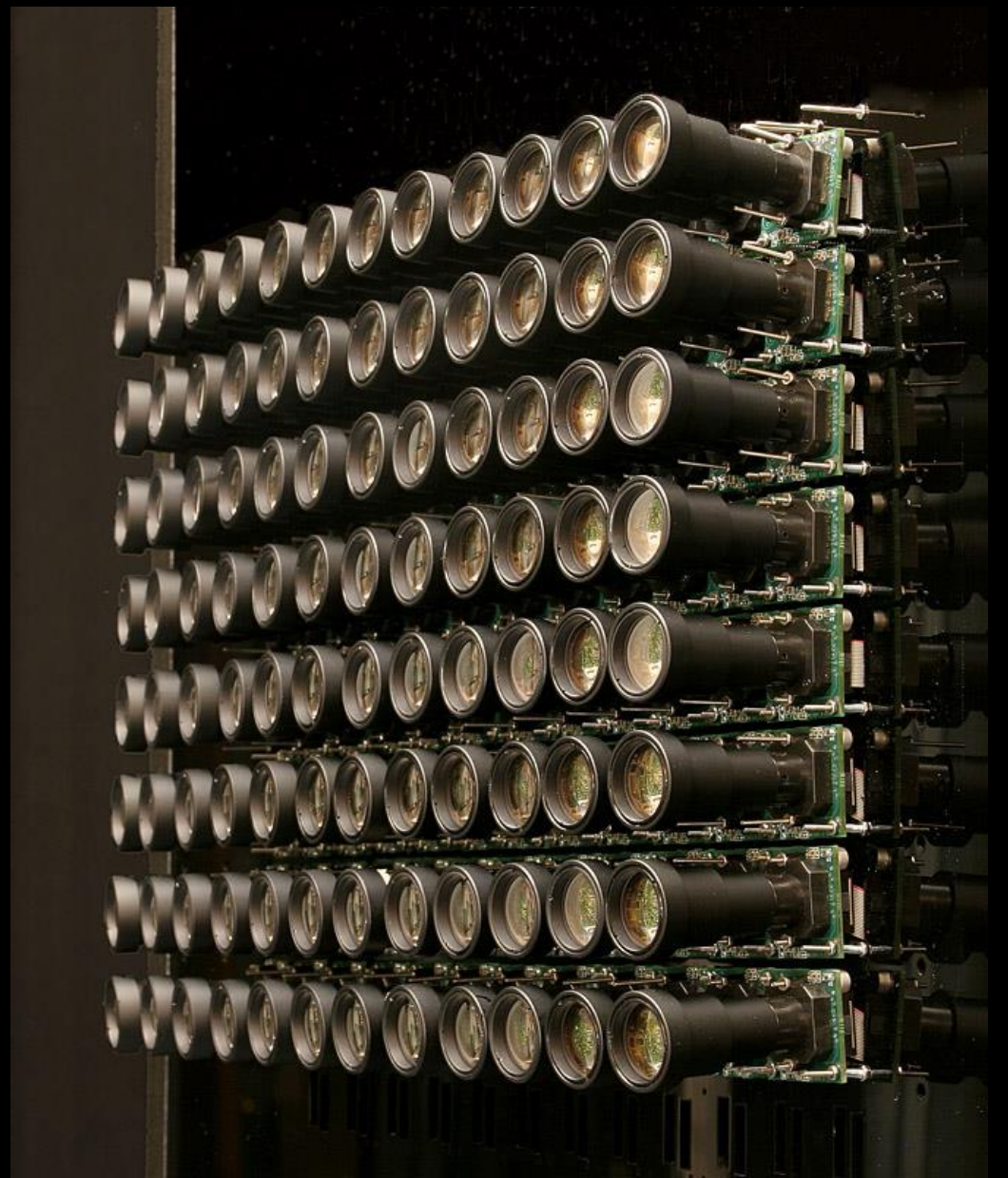
Anpei Chen

Light Fields

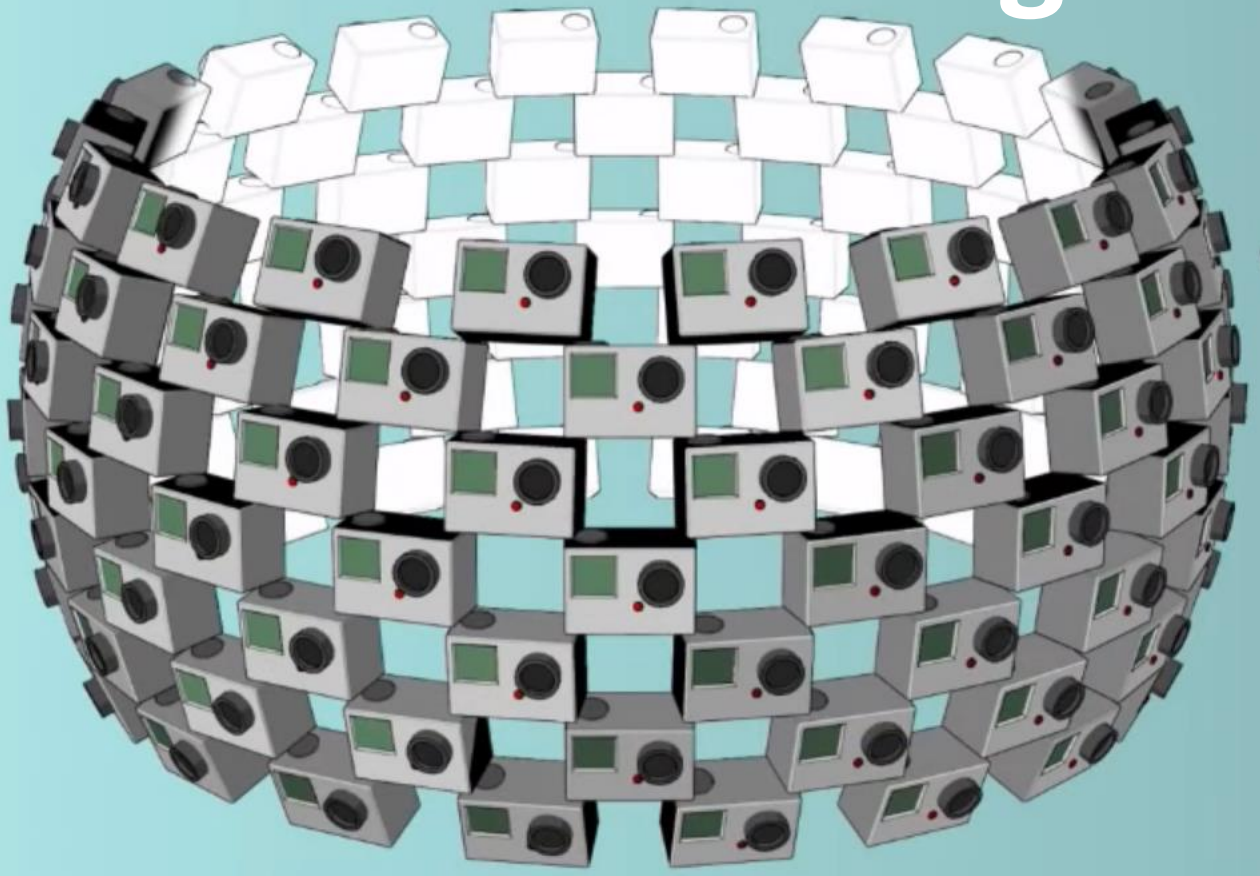
Light Fields



Plane Light field



Light Fields



18 x 7 fisheye cylindrical/spherical camera array

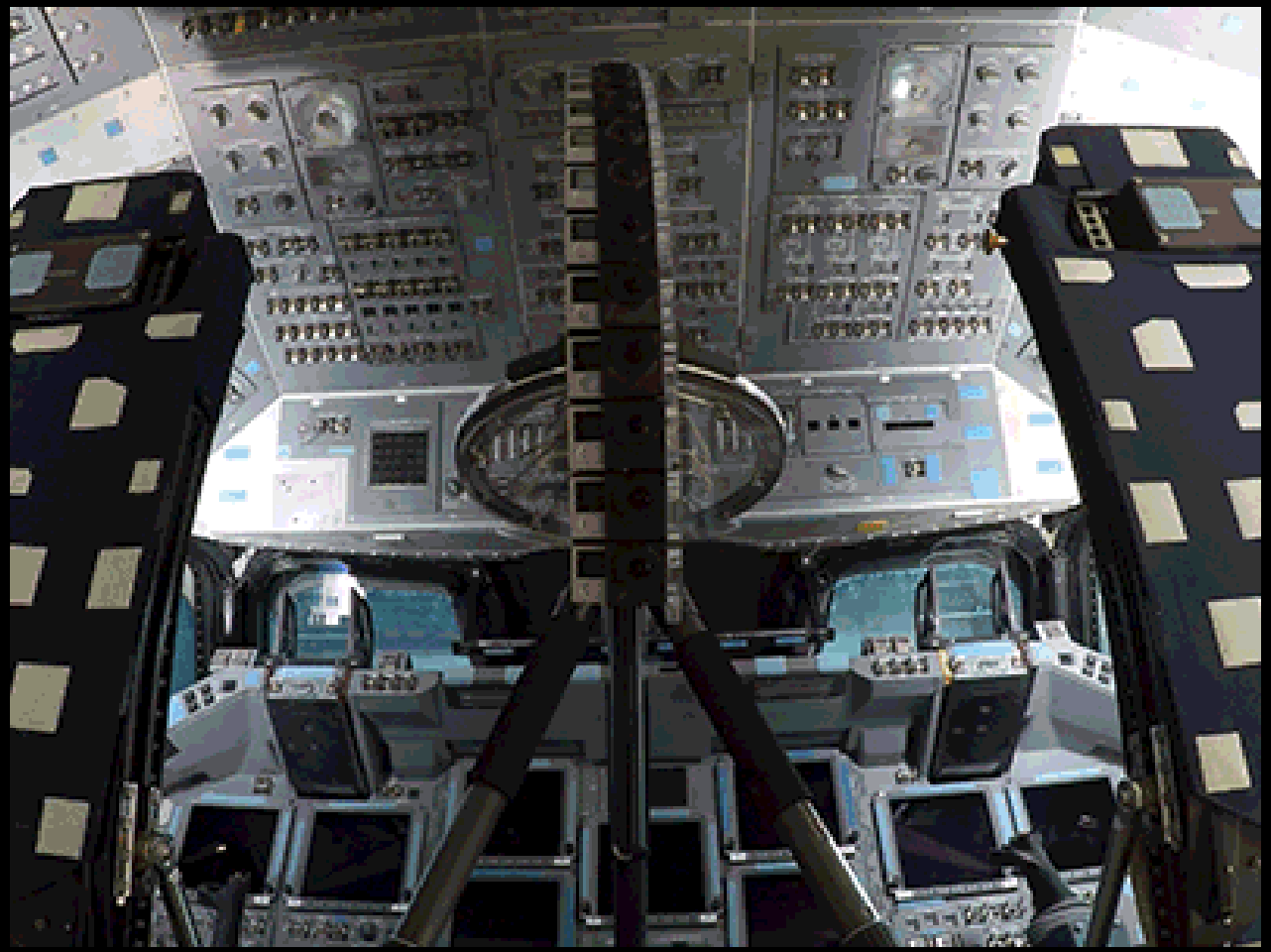
(>1Gpixel/frame at 4K each camera)

light field video playback with panoramic stereo and full parallax

USC Institute for Creative Technologies

(from Stanford SCIEN Workshop on Light Field Imaging, 2/12/2015)

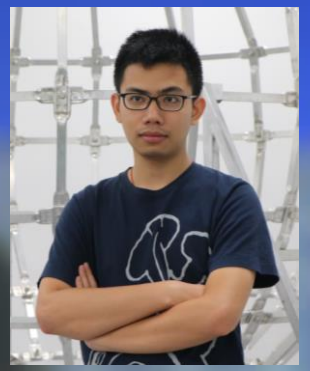
Light Fields



2018. Learn About Google’s Efforts To Capture Light Fields.

Light Fields

Deep Surface Light Fields



Anpei Chen[1]

Minye Wu[1]

Yingliang Zhang[1]

Nianyi Li[2]

Jie Lu[1]

Shenghua Gao[1]

Jingyi Yu[1,3]

[1] ShanghaiTech University [2] Duke University [3] PlexVR Inc

双11全球狂欢节
2016 11.11 GLOBAL SHOPPING FESTIVAL
23: 14: 25
¥ 116,083,052,393
$ 17,099,957,286
Total GMV in USD
初步数据，未经审计。All data on the screen are preliminary and unaudited.
无线成交占比
Mobile GMV Ratio
82%
06:54:52
571亿
¥ 57.1 Billion
15:19:13
912亿
¥ 91.2 Billion
18:55:36
1000亿
¥ 100 Billion
All GMV referenced is settled through Alipay.

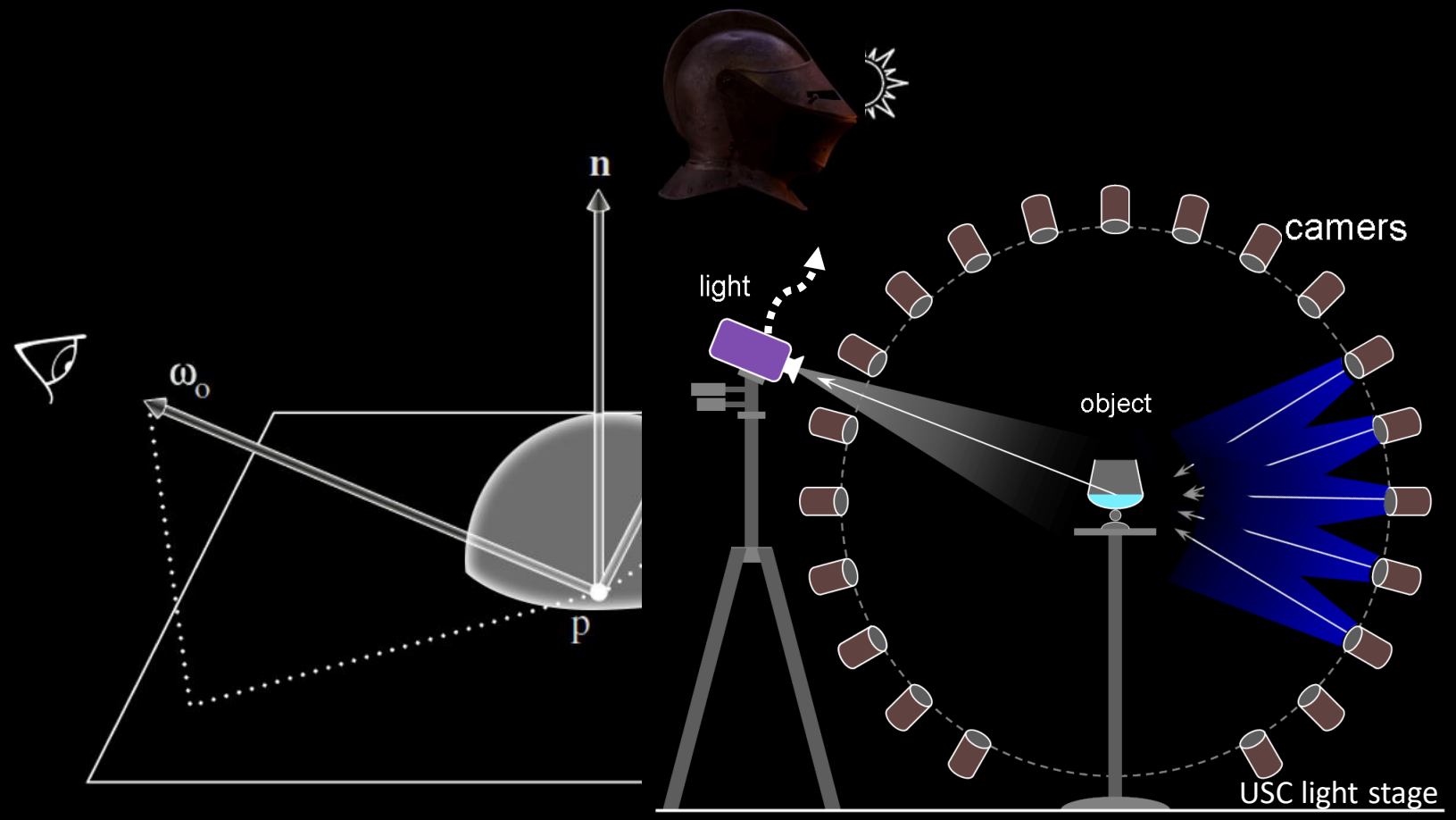
What’s Next for E-Commerce?

What’s Next for E-Commerce?

Next-Generation E-Commerce

Traditional Solutions

How to Render Reflectance?



Physically based **Image-based**

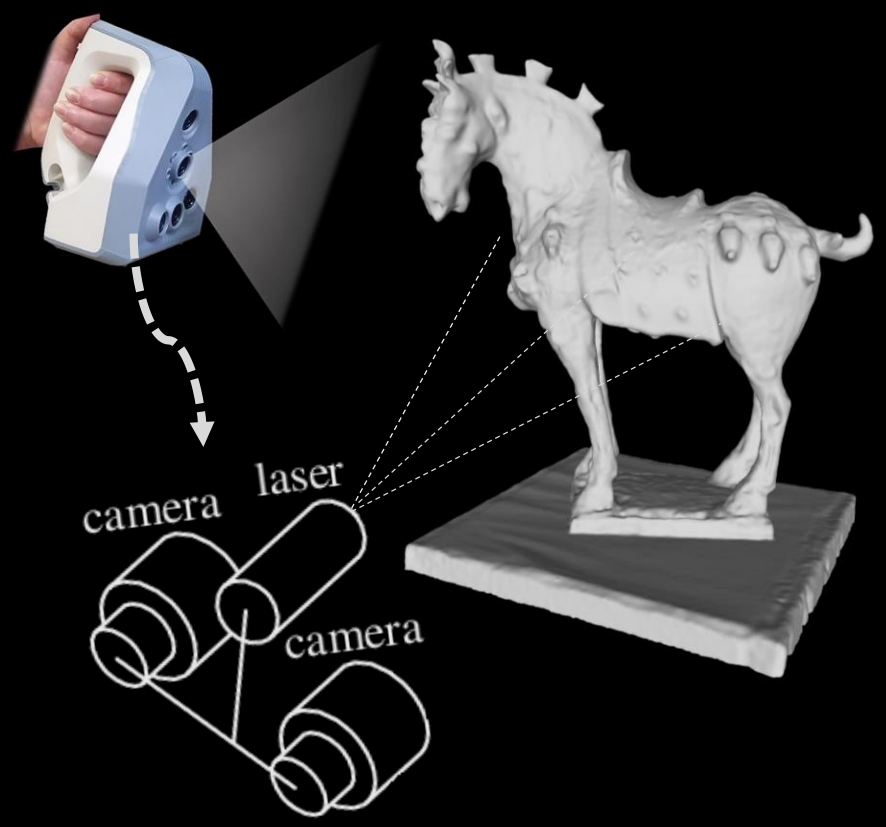
Complete Relighting

Input Image

Relit Image
(Lower-right: input target light)

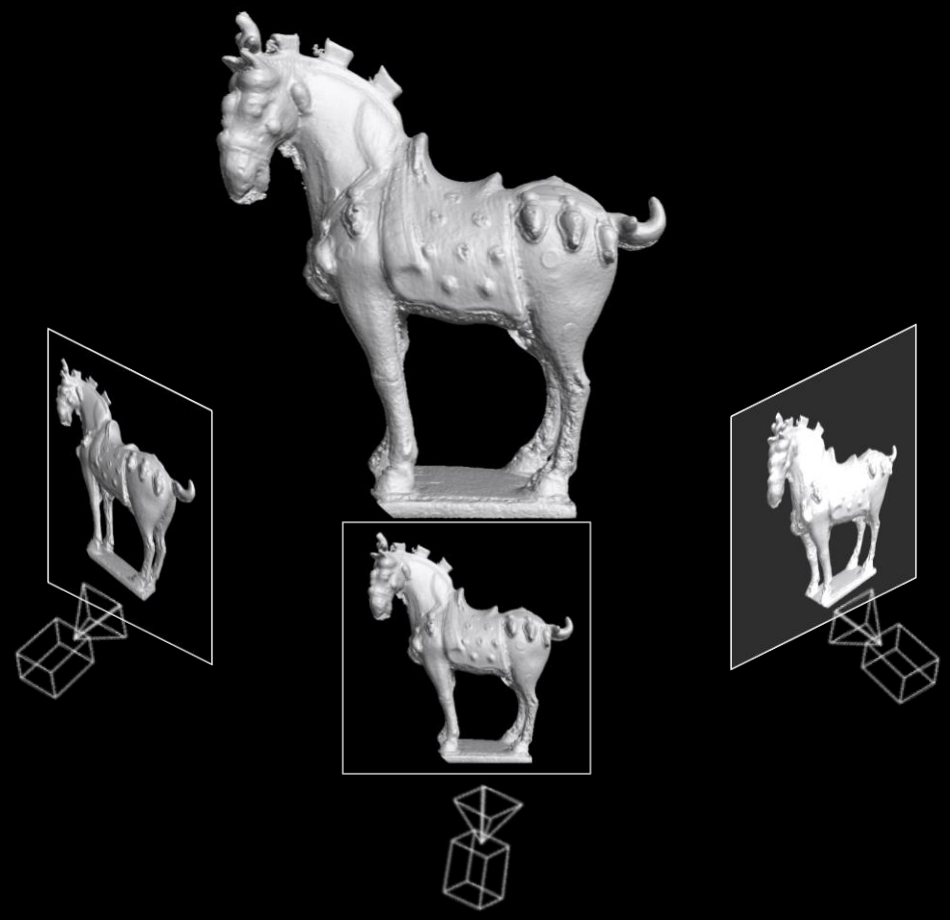
gifs.com

Using a 3D Scanner



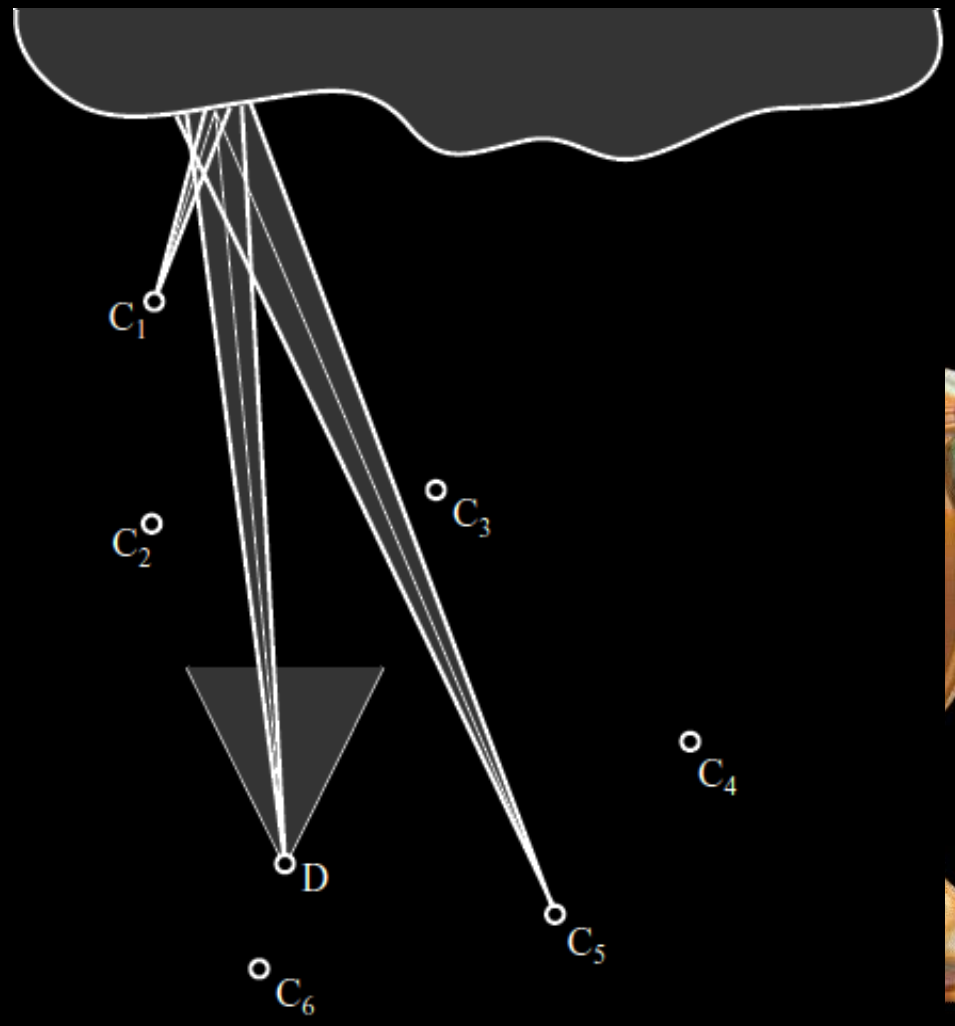
Levoy et al '2000. The Digital Michelangelo Project: 3D Scanning of Large Statues.

Multi-View Reconstruction



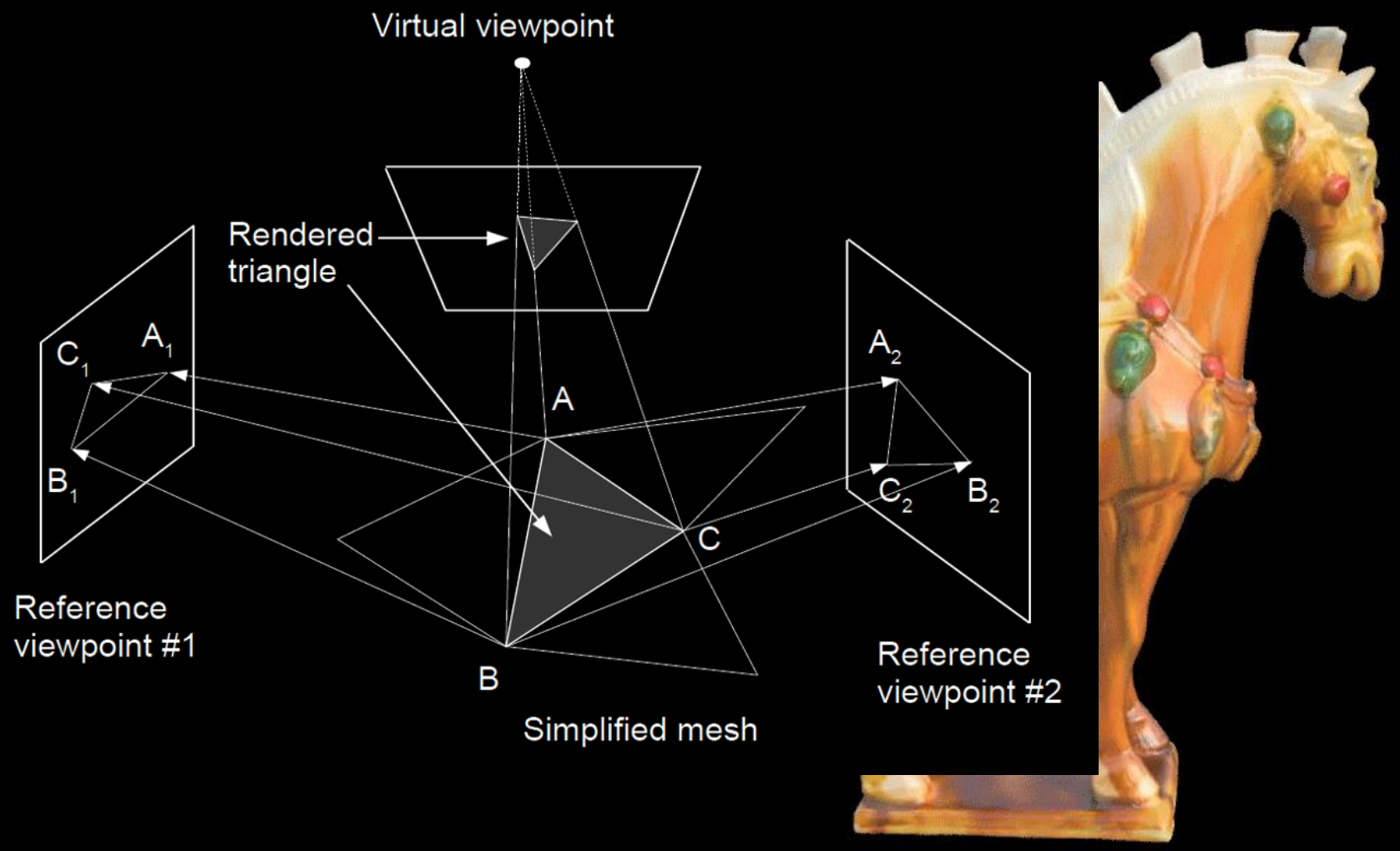
Schonberger et al. 2016. Structure-from-Motion Revisited.

Unstructured Lumigraph Rendering (ULR)



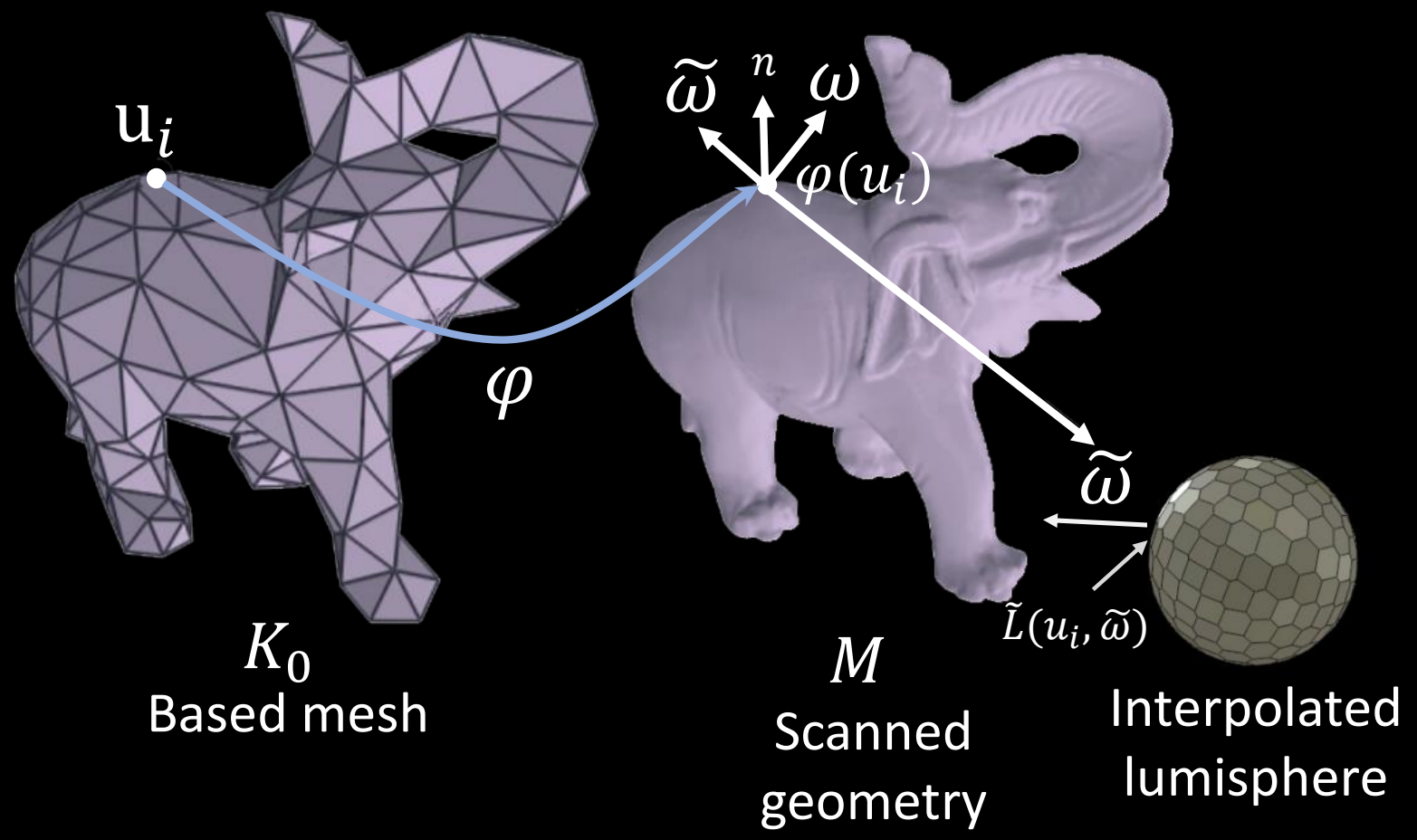
Buehler et al. 2001. Unstructured lumigraph rendering.

View Dependent Texture Mapping (VDTM)



Porquet et al. 2005. Realtime high-quality view-dependent texture mapping using per-pixel

Surface Light Fields



$$\varphi : K_0 \rightarrow m \subset \mathbb{R}^3$$

$$L : K_0 \times S^2 \rightarrow RGB$$

Daniel N Wood, et al. 2000. Surface light fields for 3D photography.

Deep Surface Light Fields (DSLF)

Learning-based Rendering

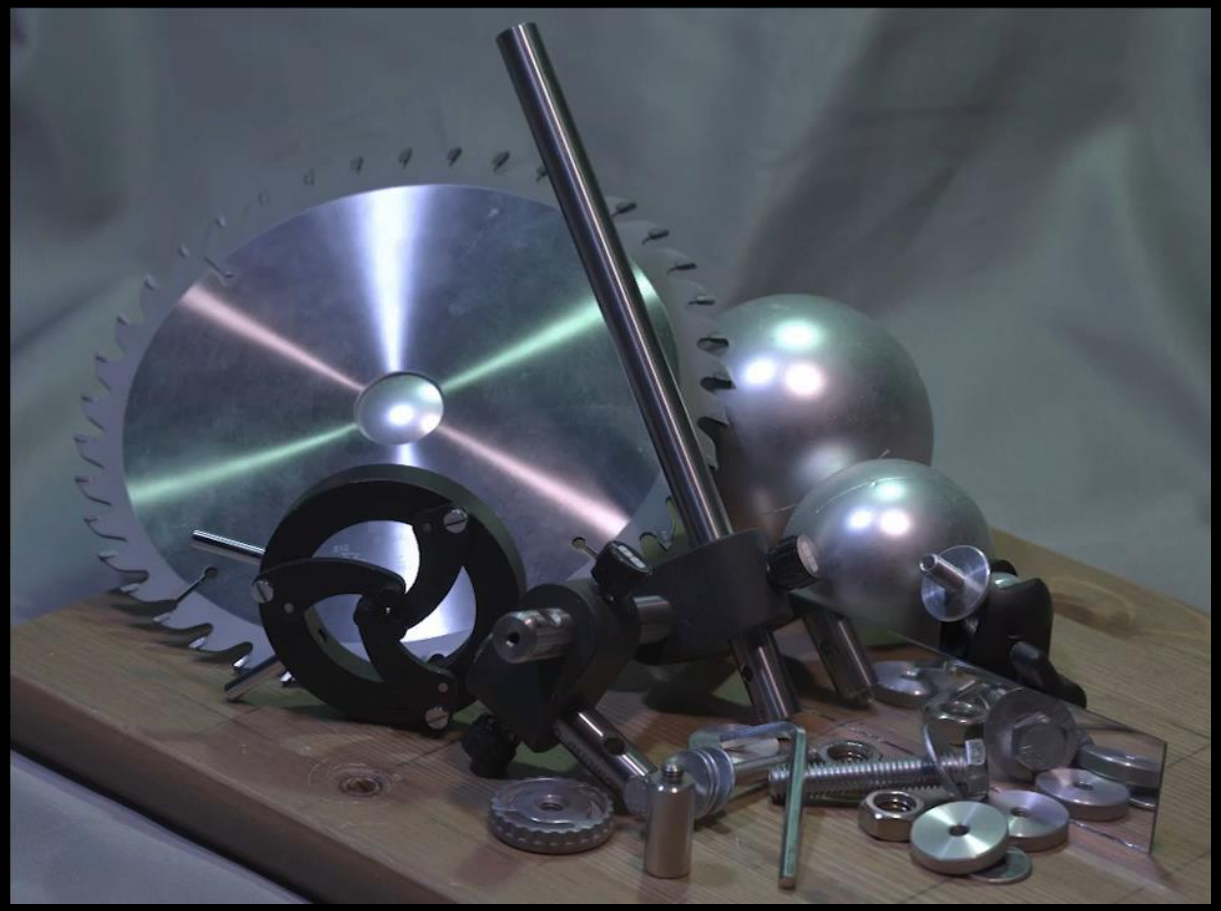


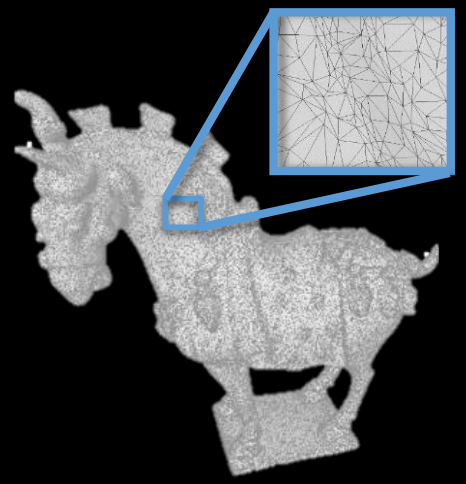
Image Based Relighting

Global Illumination Prediction

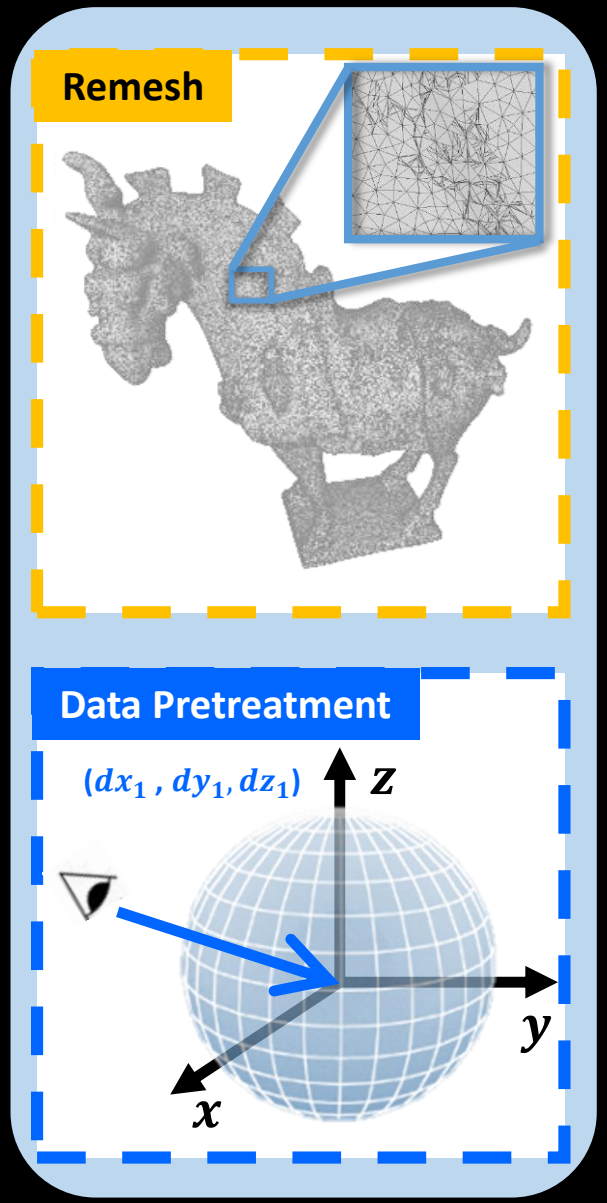
Ren et al. 2015. Image based relighting using neural networks.
Ren et al. 2013. Global illumination with radiance regression functions.

Algorithm Overview

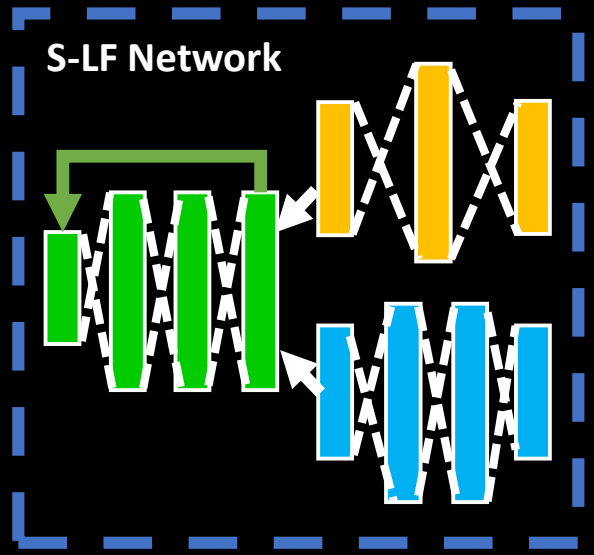
Images registration



Base mesh



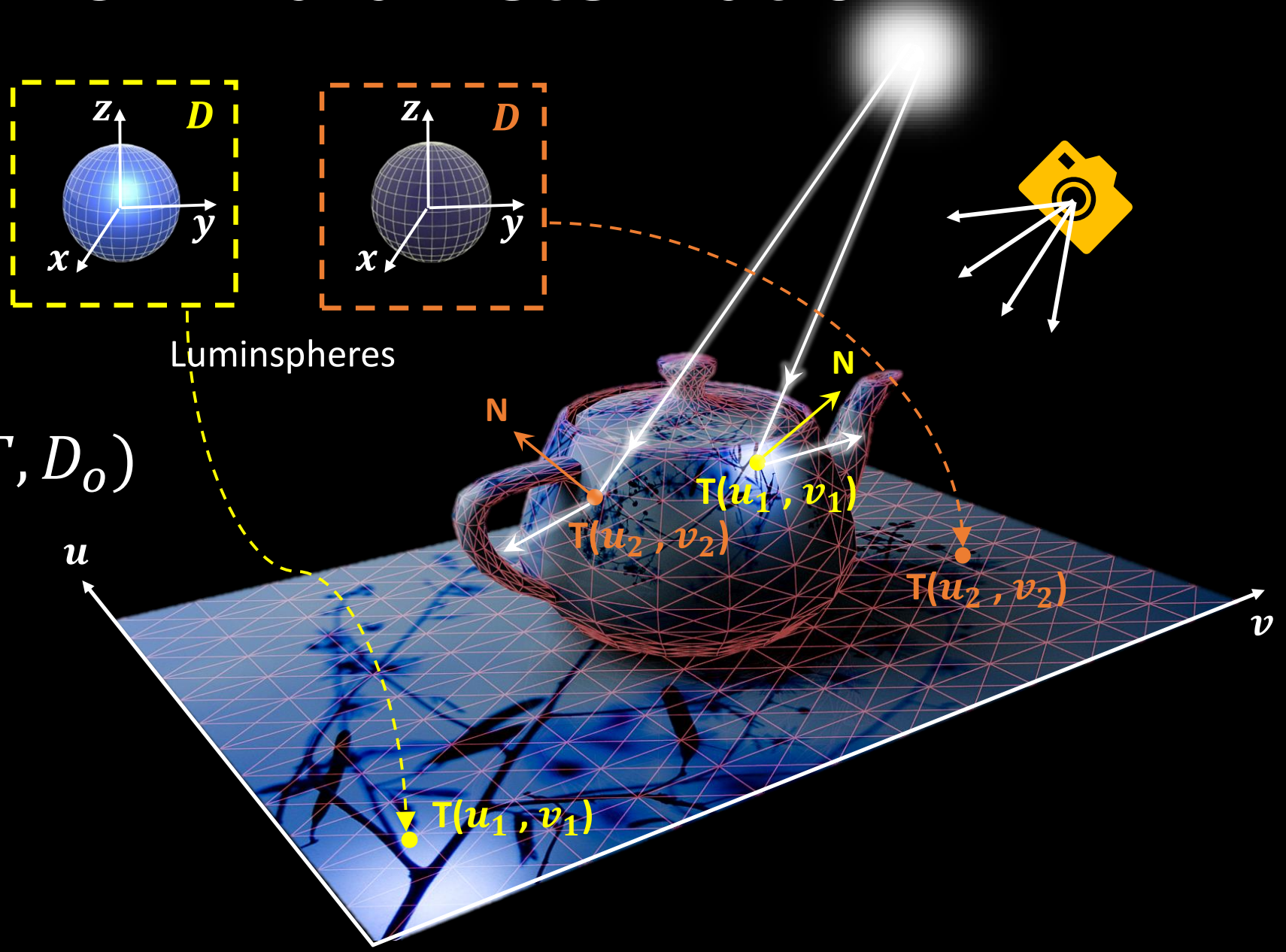
Data pre-processing



Training

DSLF Parameterization

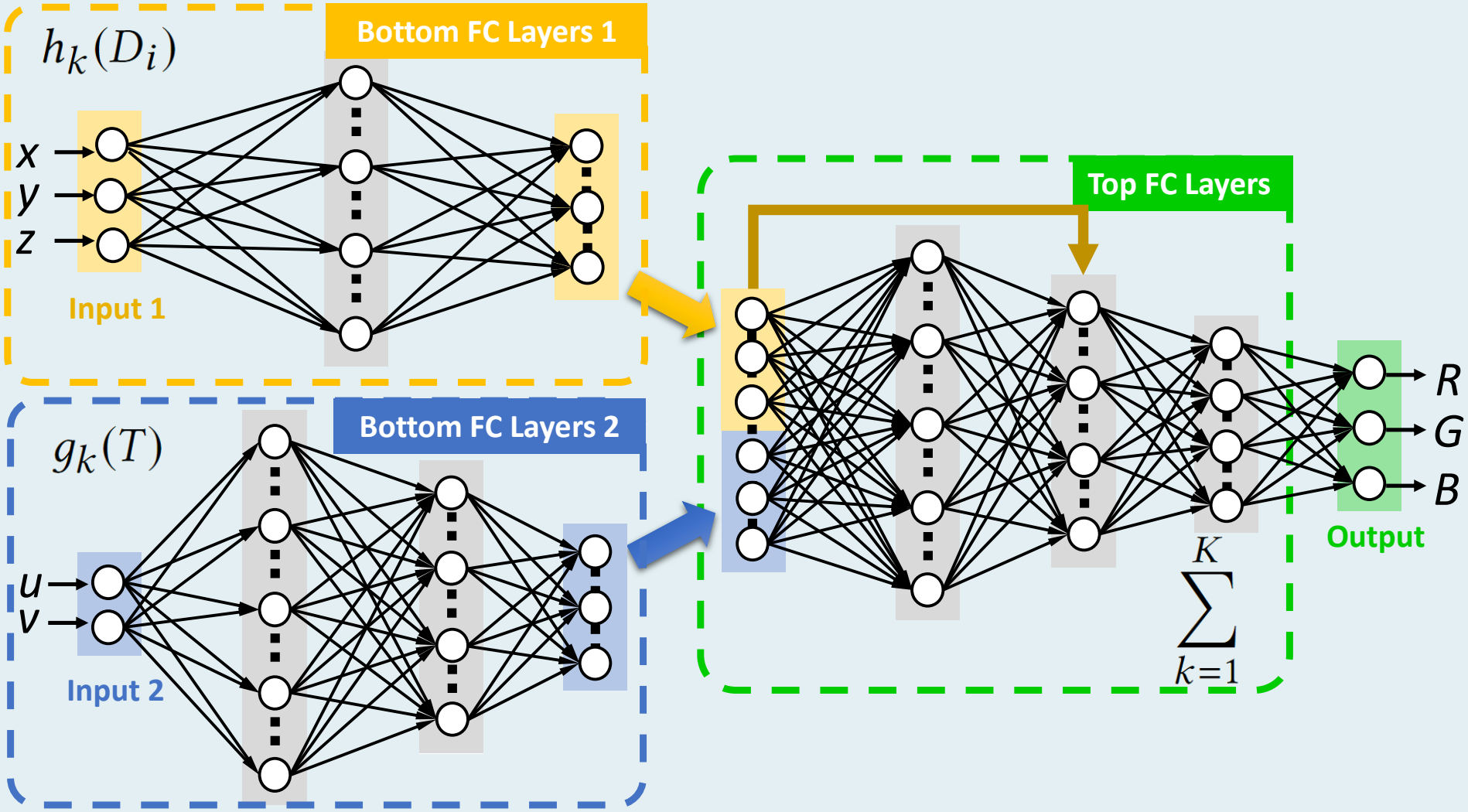
$$S(T, D_o) \approx \Phi_d(T) + \Phi_r(T, D_o)$$



Our DSLF network

Architecture

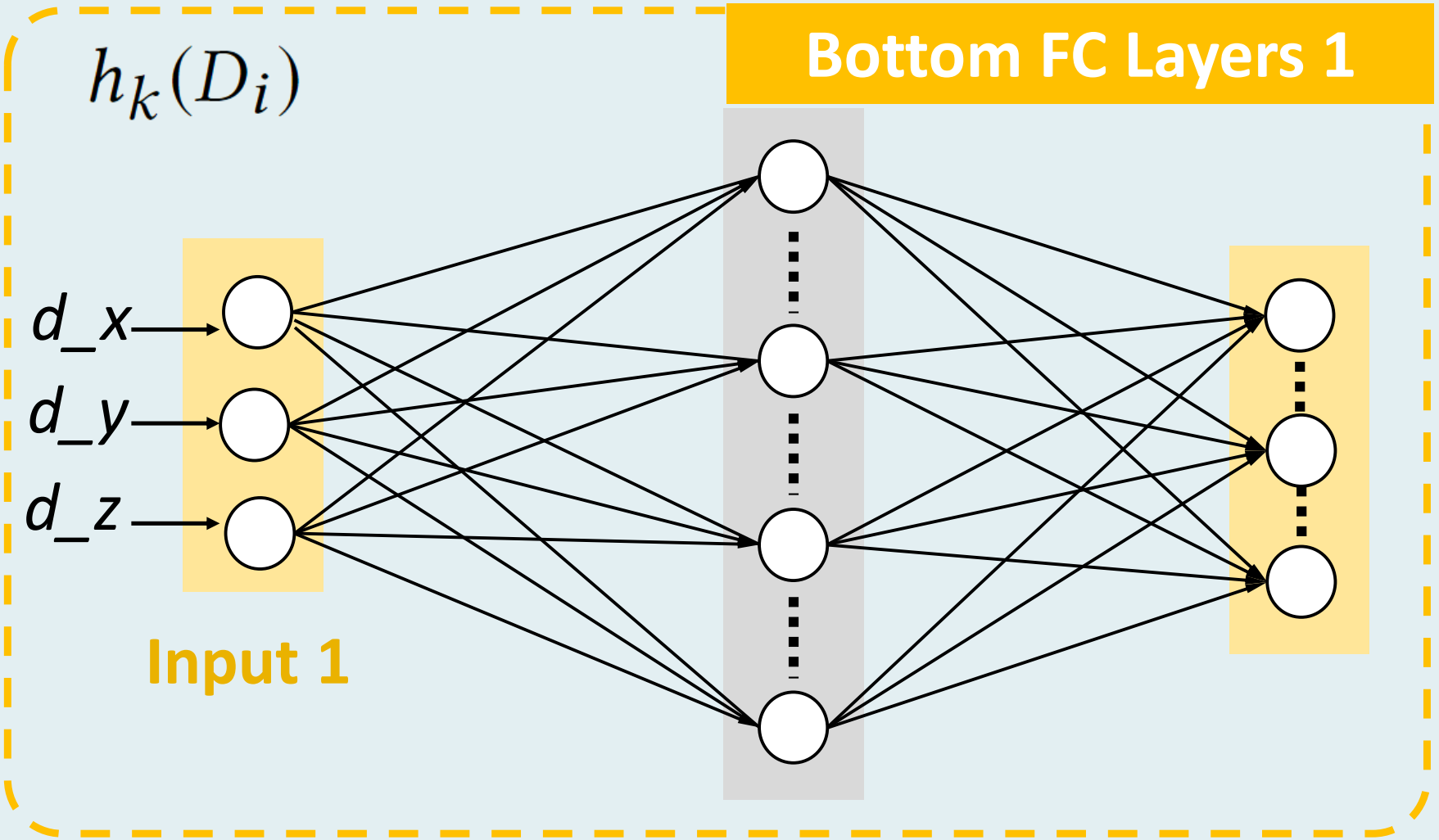
$$\Phi_r(T, D_i) \approx \sum_{k=1}^{K} g_k(T)\, h_k(D_i)$$



Directional Component

$$\Phi_r(T, D_i)$$

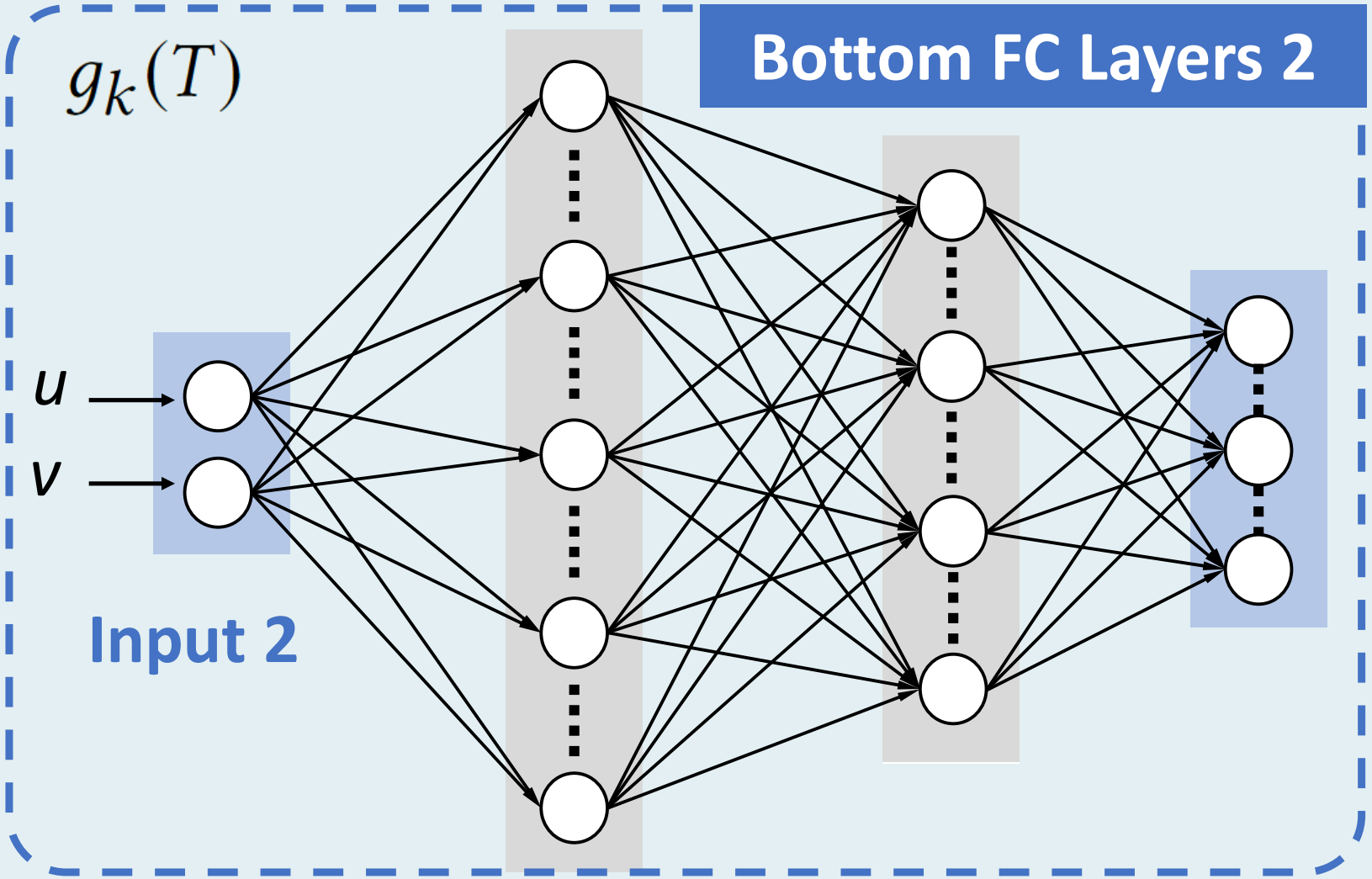
$$\approx \sum_{k=1}^{K} g_k(T)\, h_k(D_i)$$



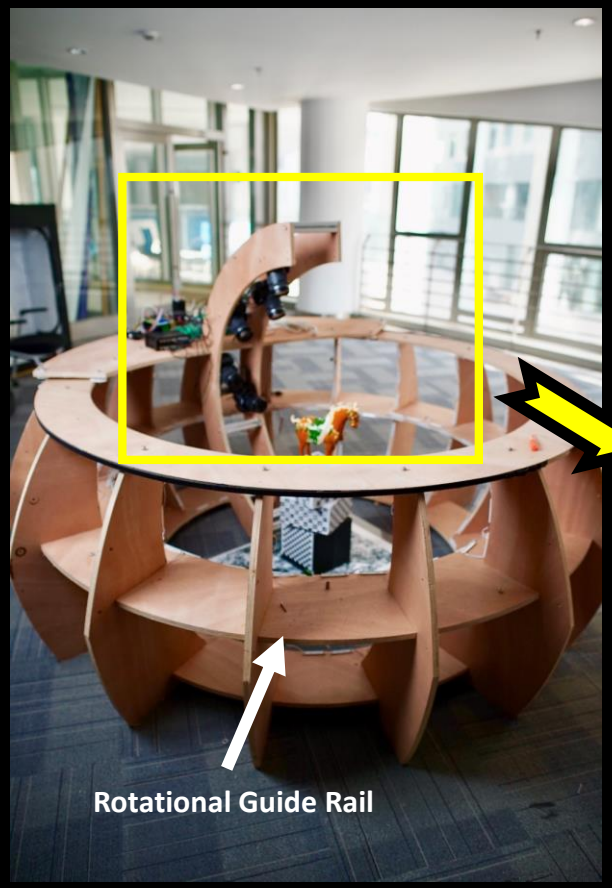
Vertex Component

$$\Phi_r(T, D_i)$$

$$\approx \sum_{k=1}^{K} g_k(T)\, h_k(D_i)$$

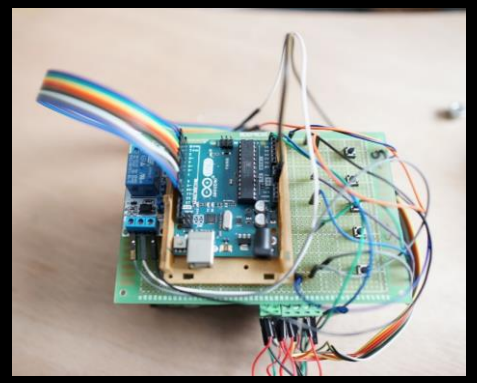


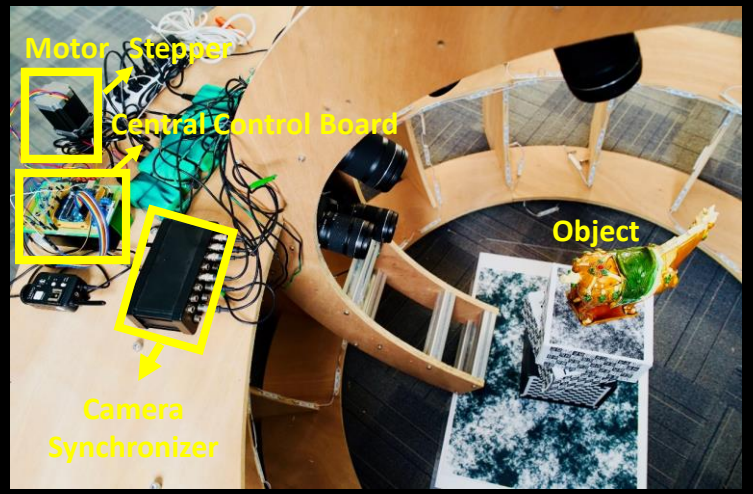
Data Aacquisition and Registration



Motor Stepper

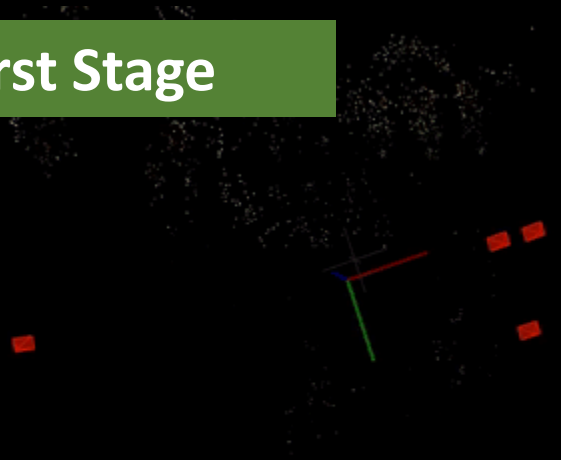
Central Control Board





Capture Device

First Stage

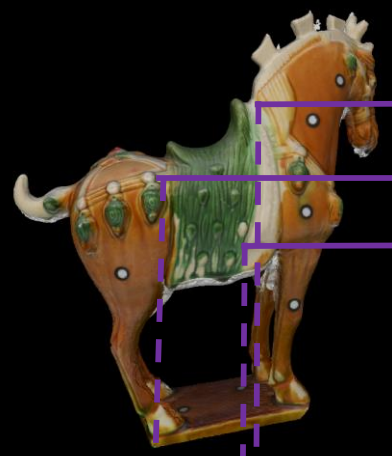


Second Stage

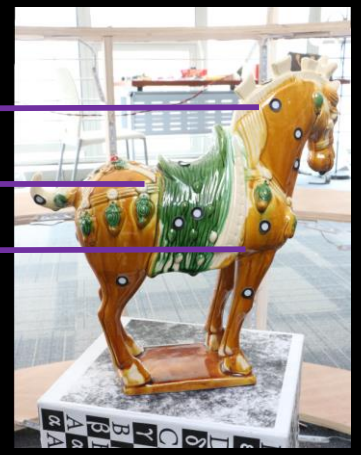
Registration

Image Registration

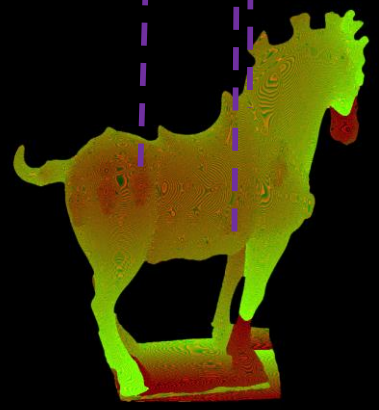
3D Model from First Stage



Color Image



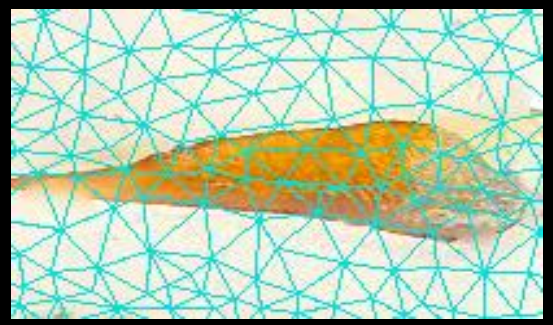
Captured Image



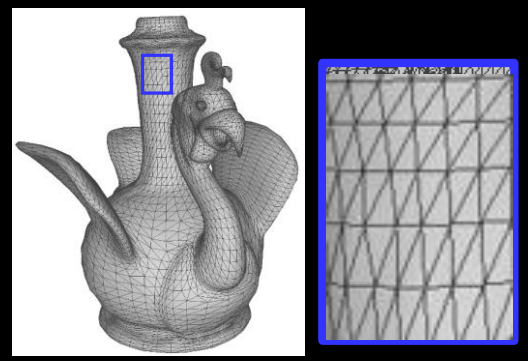
Depth Image

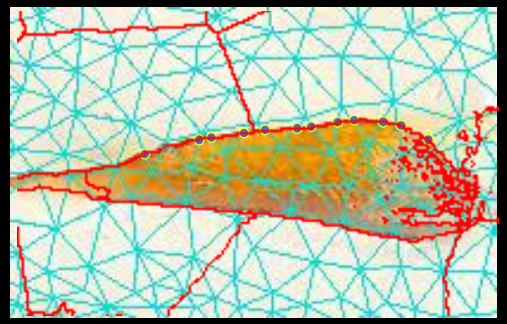
Solve as a Perspective-n-Point Problem

Texture-Aware Remeshing

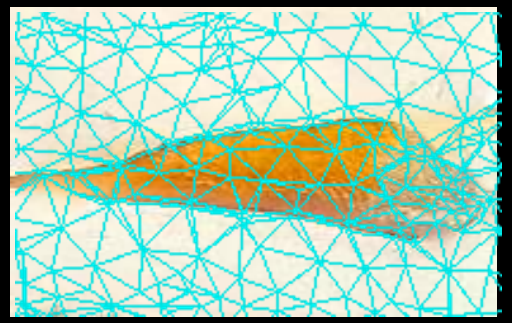


Base mesh

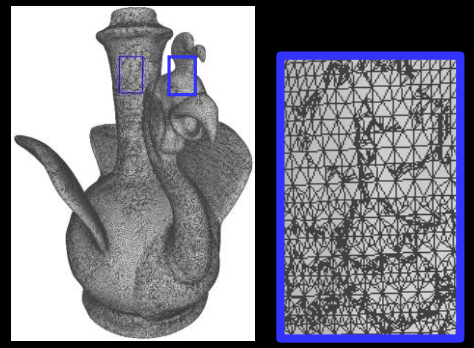


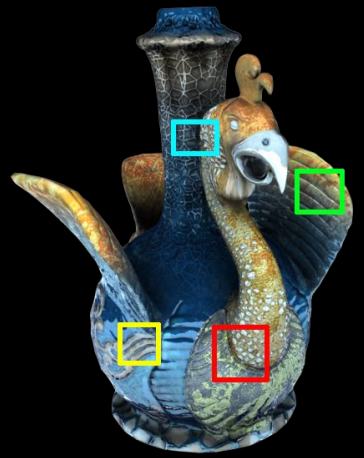


Superpixel



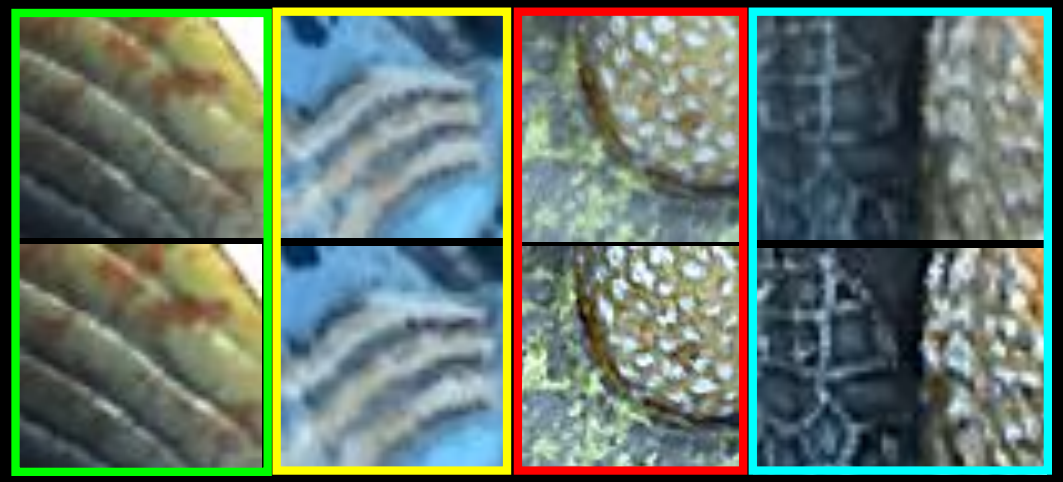
Remeshing





Before

After



Texture-Aware Remeshing

What is a good triangle?

- **Monotonic**, the gradients are consistently positive or negative
- **Smooth**, the second order derivative is nearly zero

Energy function:

$$E = C + \gamma G + \beta B$$

Where:

$$C = \left(\frac{1}{Z}\sum_{k}\sum_{\mathcal{A}} sign(\Delta(i))\right)^2$$

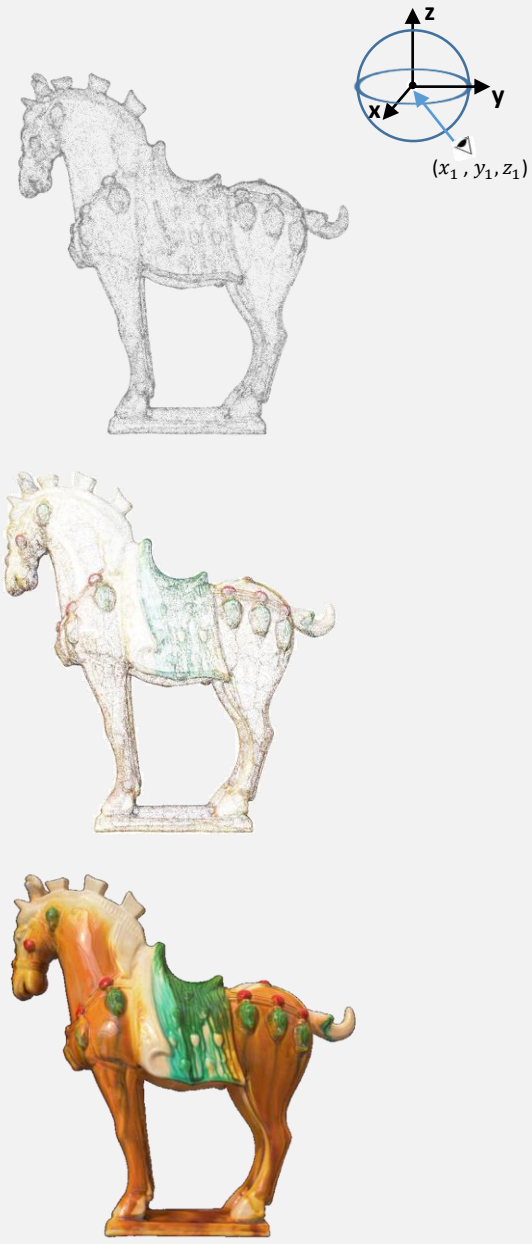
$$G = \frac{1}{Z}\sum_{k}\sum_{\{H_j\}}\sum_{i} \delta(\Phi(i) \in H_j)$$

$$B = \sum_{i}\sum_{k}\left(\frac{1}{Z}\sum_{j \in \mathcal{N}_i} \delta(j \in \mathcal{A})\right)^2$$

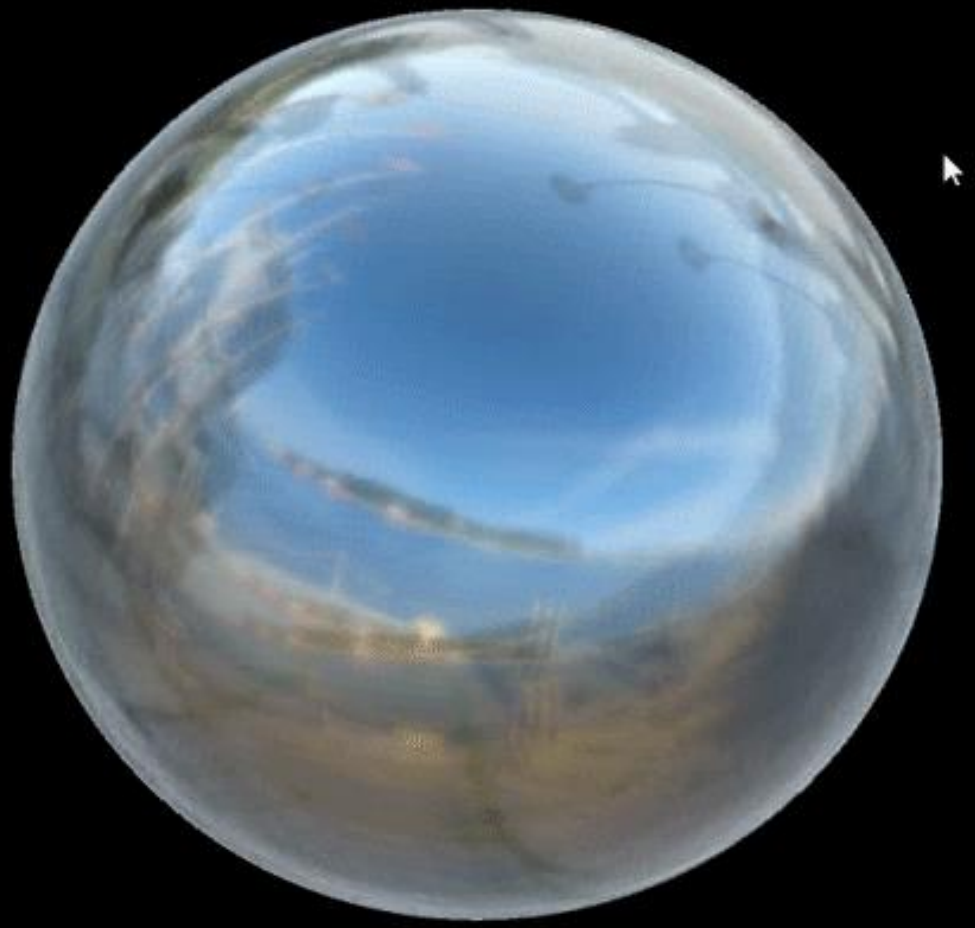
Rendering Algorithm

Algorithm 1 Rendering DSLF

```
procedure RenderOneFrame(T, N, F, C)
    (T′, N′, F′) = BackFaceCulling(T, N, F, C)
    for each vertex T_i in T′ do
        d_i = CalculateViewingDirection(T_i, C)
        d̃_i = InvertTransform(d_i, n_i)
    D̃ = (d̃_1, ..., d̃_i)
    VertexColor = DSLF(T′, D̃)
    Image = ProjectAndRasterizate(C, T′, F′, VertexColor)
    return Image
```

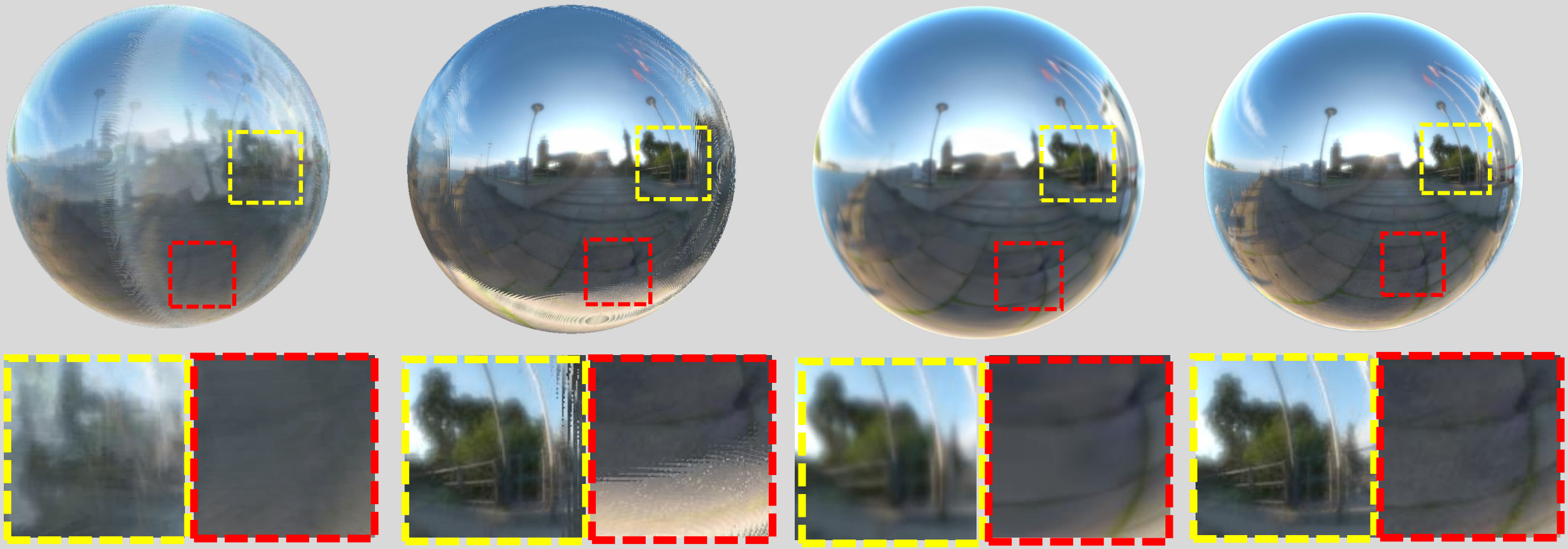


Synthetic: Metal Sphere



Vertex #	521,962
Image #	200
Resolution	1536*1167
Superpixel #	6,000

Closeup Comparisons



VDTM ULR Ours Photograph

Real Example: Beer Can

DSLF

Vertex #	**467,418**	**Resolution**	**1920*1280**
Image #	320	SuperPixel	4,000

Closeup Comparisons

VDTM ULR Ours Photograph

Real Example: Liquor Bottle

Vertex #	483,056
Image #	217
Resolution	1536*1167
Superpixel #	4,000

Real Example: Tang San Cai

Vertex #	441,412
Image #	216
Resolution	1536*1167
Superpixel #	4,000

Rendering Quality

PSNR	DSLF	VDTM	ULR
Metal Sphere	27.1997	19.0778	21.6797
Synthetic Tang San Cai	28.5230	23.6546	25.4801
Beer Can	18.0536	13.0823	14.5006
Liquor Bottle	22.2871	20.8831	20.8893
Real Tang San Cai	**20.4035**	**21.4786**	**21.6177**

Rendering Quality

SSIM	DSLF	VDTM	ULR
Metal Sphere	0.9194	0.8164	0.8339
Synthetic Tang San Cai	0.9687	0.9415	0.9594
Beer Can	0.6555	0.4851	0.5741
Liquor Bottle	0.8099	0.7727	0.7912
Real Tang San Cai	0.9480	0.9242	0.9350

Compression Rate

	Metal Sphere	Tang San Cai	Beer Can	Liquor Bottle	Tang San Cai
Image size(GB)	2.003	2.15	4.89	3.36	3.35
Train Time(hour)	1.1	1.3	3.0	2.3	1.6
Compressed Output(MB)	0.799	0.799	0.799	0.799	0.799
Compression Rate	2567:1	2760:1	6267:1	4309:1	4289:1
Rendering Speed(FPS)	42	41	38	38	35

Synthetic Real

Conclusions and Future Work

Summary

- A novel deep network based surface light field modeling and rendering technique.
- High quality rendering results
- High compression rate

Limitation

- Long training time
- Sensitive to self-occlusions

Future Work

- Extend DSLF on mobile GPUs.
- Transfer lighting to different objects

Acknowledgement

- Anonymous I3D reviewers
- The program chairs for help on journal revisions
- Shanghai Grant 17XD1402900, 17JC1403800
- Alibaba for providing the Tang San Cai on Taobao

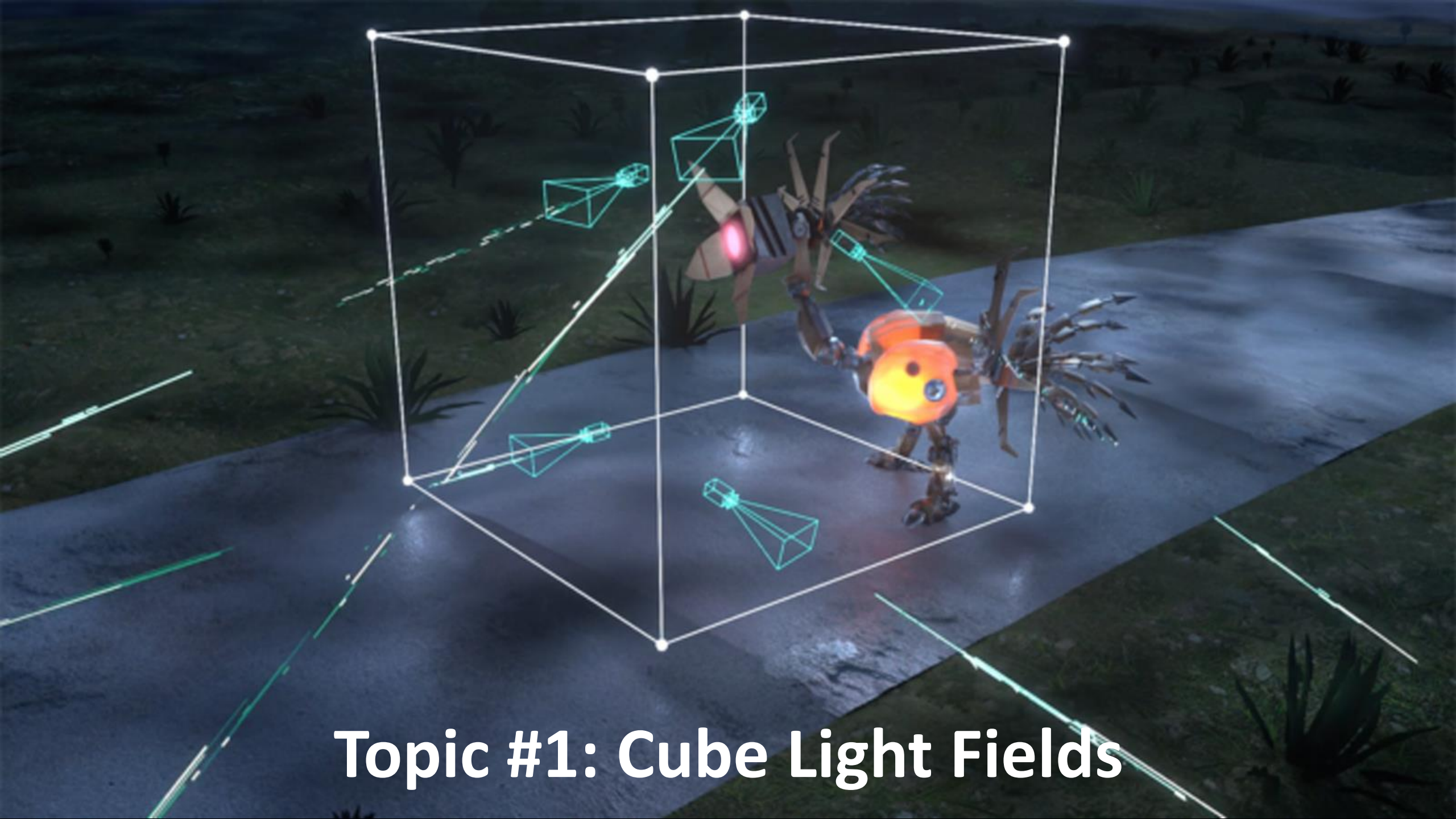
Limitations

Hard to capture data
Lot of redundancy
Hard to render high quality images
Limited on application

What’s Next?

Rendering is Magic

~~Hard to capture data~~
~~Lot of redundancy~~
~~Hard to render high quality images~~
~~Limited on application~~

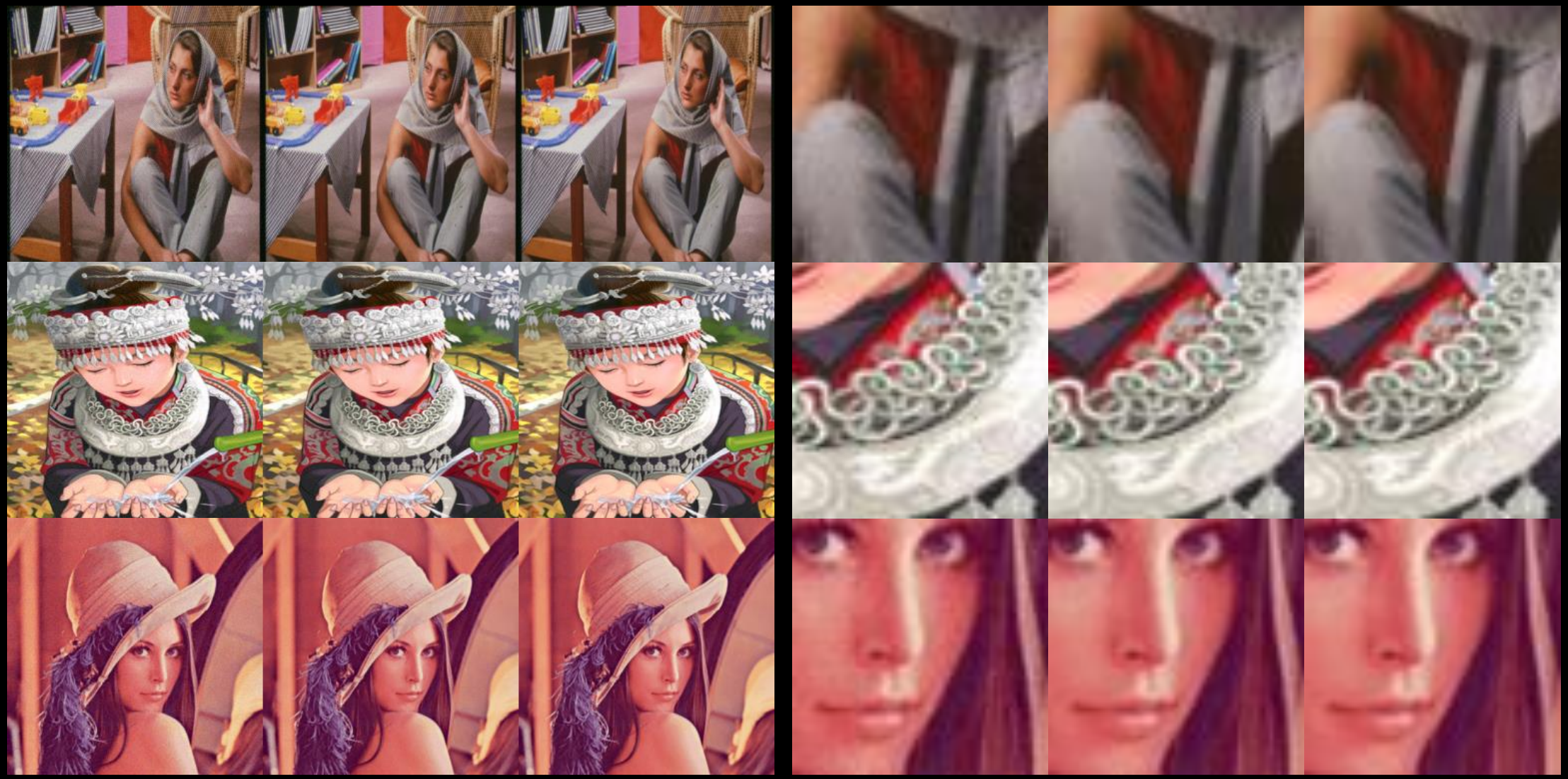


Topic #1: Cube Light Fields

Topic #2: animation rendering

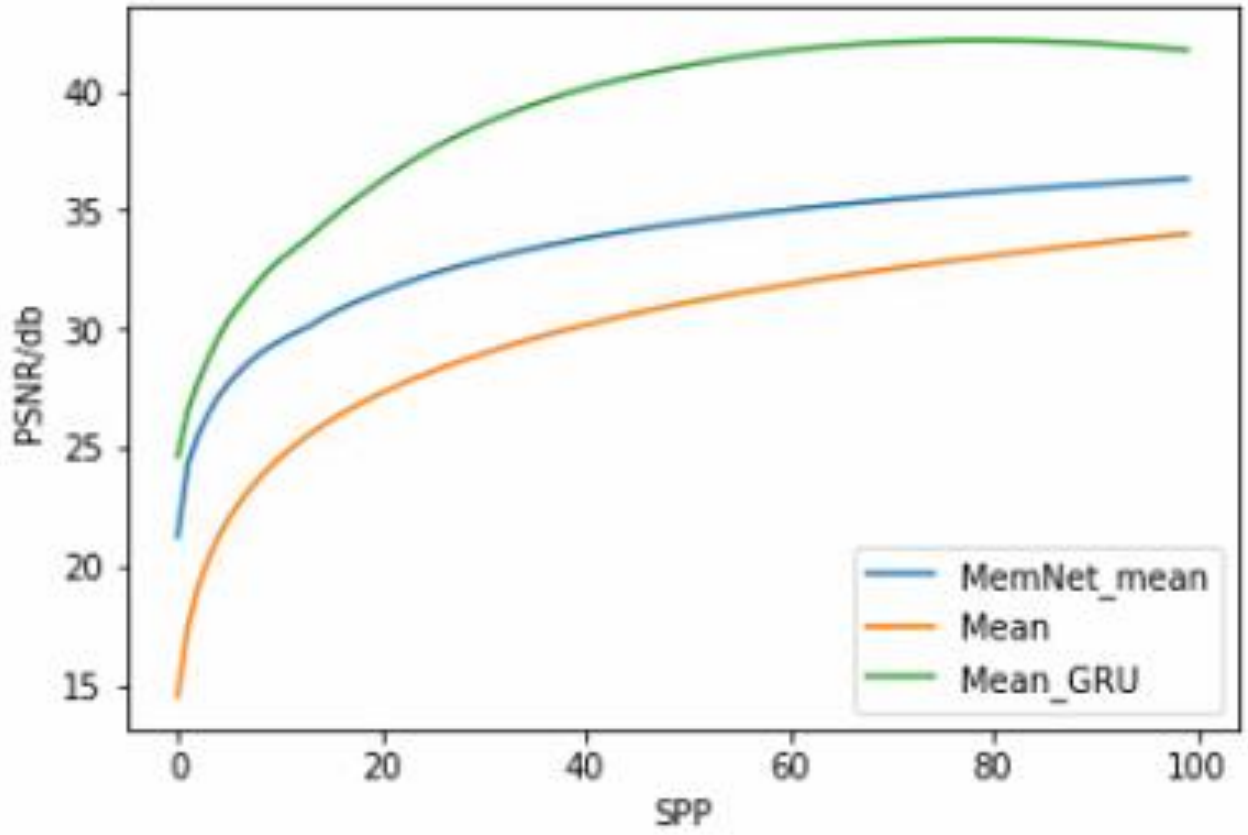
Topic #2: animation rendering

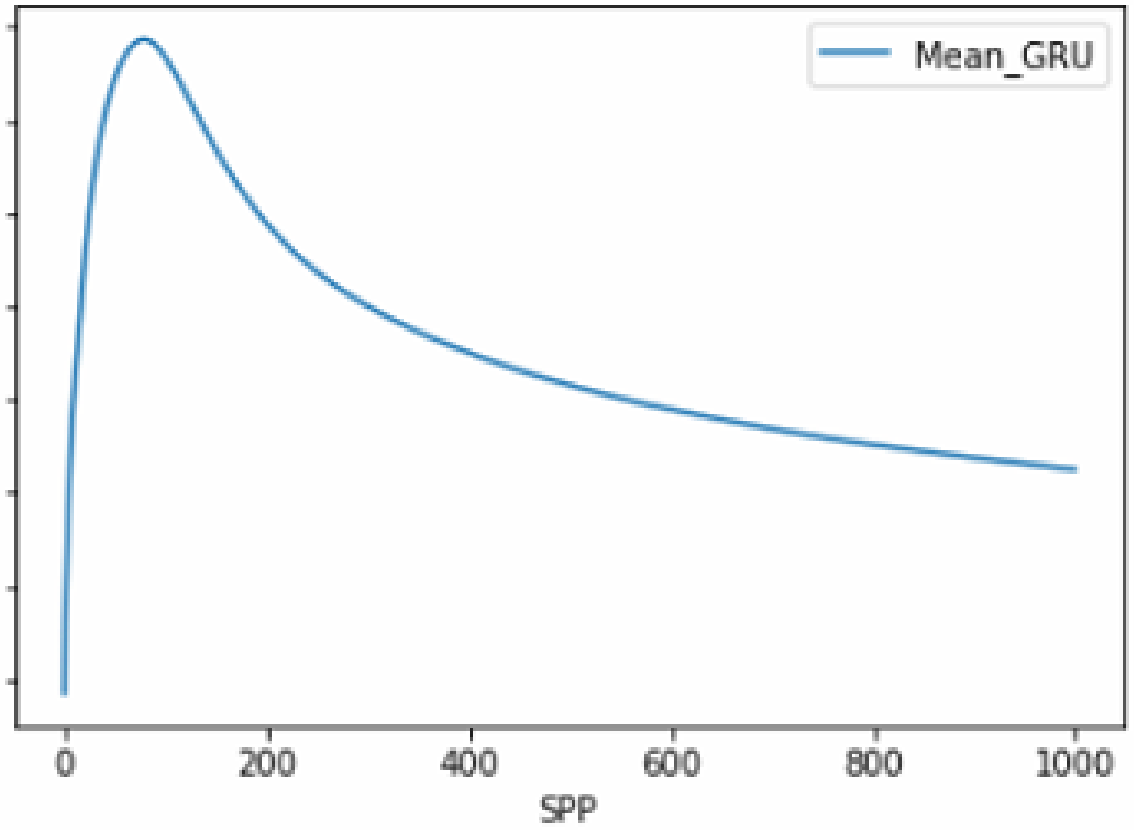
Topic #2: Experiment



Integration RNN Reference Integration RNN Reference

Topic #2: Experiment





a) The below figure showing the image our method performance between 1 to 100 spp, our(MemNet_GRU) out performance about 6db than only using image space denoise(MenNet). Note: the test result is a gaussian noise which std=0.7, image data were normalized to [-1,1]

b) The network only train first 50 SPP, but if test under large spp(>80) the PSNR decline

Topic #3: Free Viewpoint in the Wild

Topic #3: Free Viewpoint with Animation