SIGGRAPH 2021
VIRTUAL 9-13 AUGUST

Free-form Scanning of Non-planar Appearance with Neural Trace Photography





Xiaohe Ma[1] Kaizhang Kang[1] Ruisheng Zhu[1] Hongzhi Wu[1] Kun Zhou[1,2]

[1] State Key Lab of CAD&CG, Zhejiang University
[2] ZJU-FaceUnity Joint Lab of Intelligent Graphics

Introduction

Introduction

- Realistic Material Appearance is Important

Culture Heritage

e-Commerce

Visual Effects

Lycurgus Cup
©The British Museum

https://www.amazon.cn/dp/B0041IXBGW/ref=sr_1_102?dchild=1&qid=1621213201&s=luggage&sr=1-102

© Paramount Pictures

Introduction

- Capturing Appearance is Challenging

3D Scanner

Multiview Stereo

3D Mesh

+

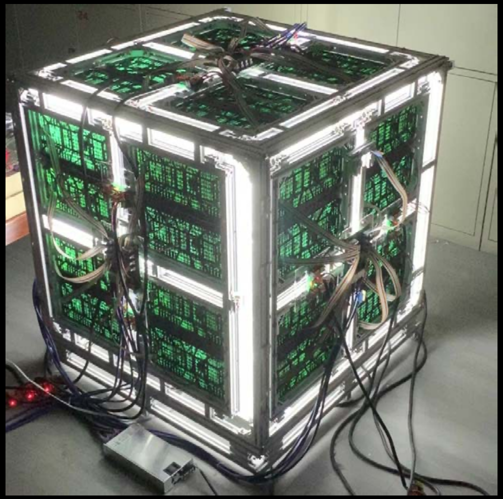
6D SVBRDF
(Varies with Location, Lighting & View)

→

Digital Model

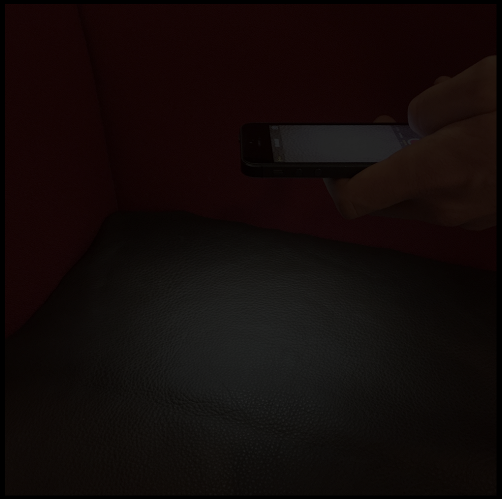
Introduction

Lightstage



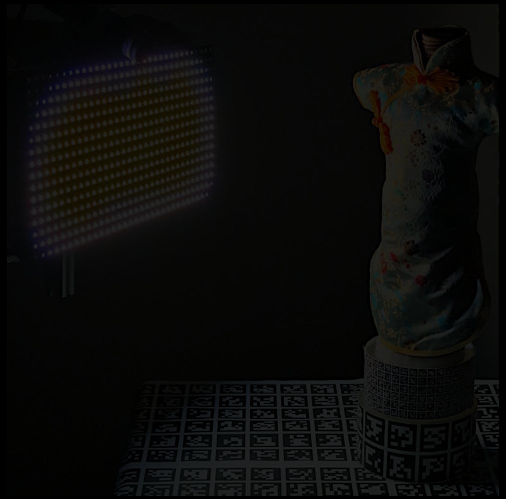
[Kang et al. 2019]

Camera-Flash Pair



[Aittala et al. 2015]

Our Goal



	Lightstage	Camera-Flash Pair	Our Goal
Sampling Efficiency	High	Low	High
Spatial Coherence	No	Yes	No
Anisotropic	Yes	No	Yes
Movable	No	Yes	Yes
Max Sample Size	Limited	Unlimited	Unlimited

Differentiable Framework

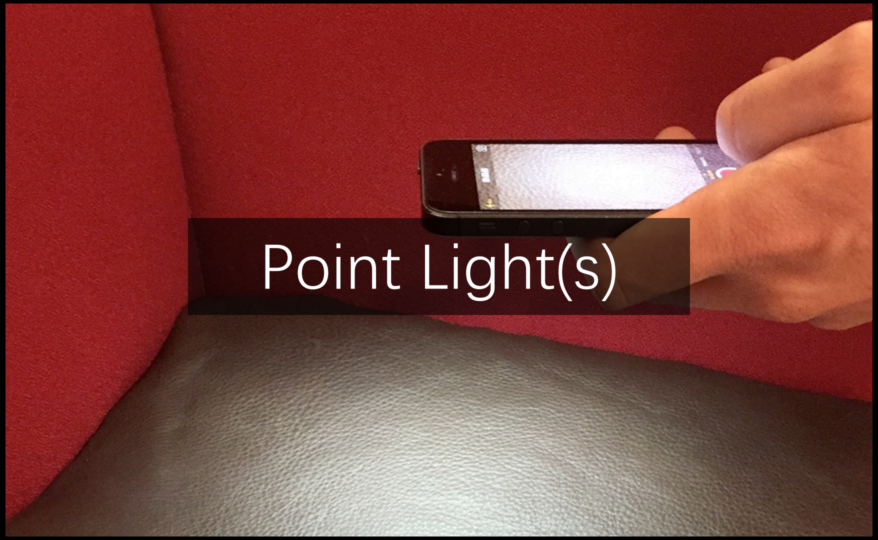
- High-quality Scanning of Anisotropic Appearance
- Automatically Learns
  - Lighting Condition
  - Measurements => Reflectance
- Adapts to Various Factors
  - Point/Linear/Area Light
  - Setup's Geometry

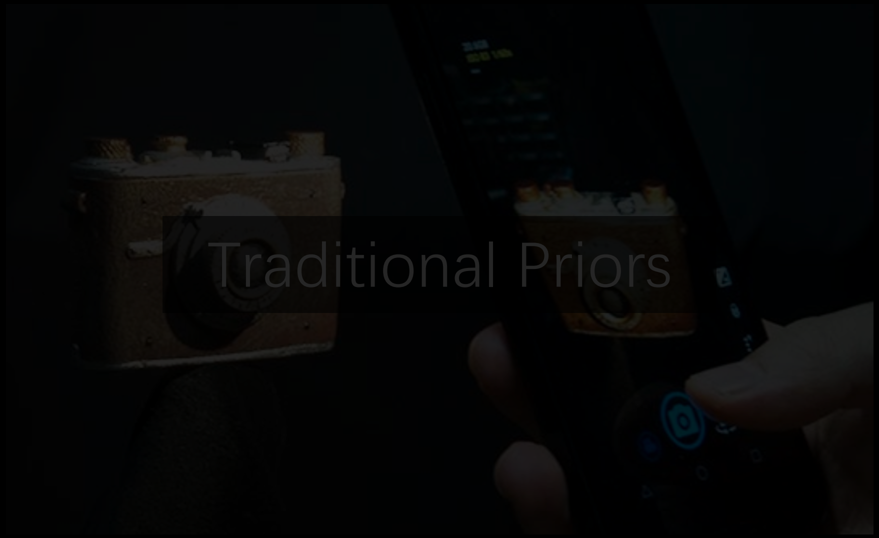
Our Scanned Results

Key Insight: Appearance Scanning = Geometry Learning

Related Work

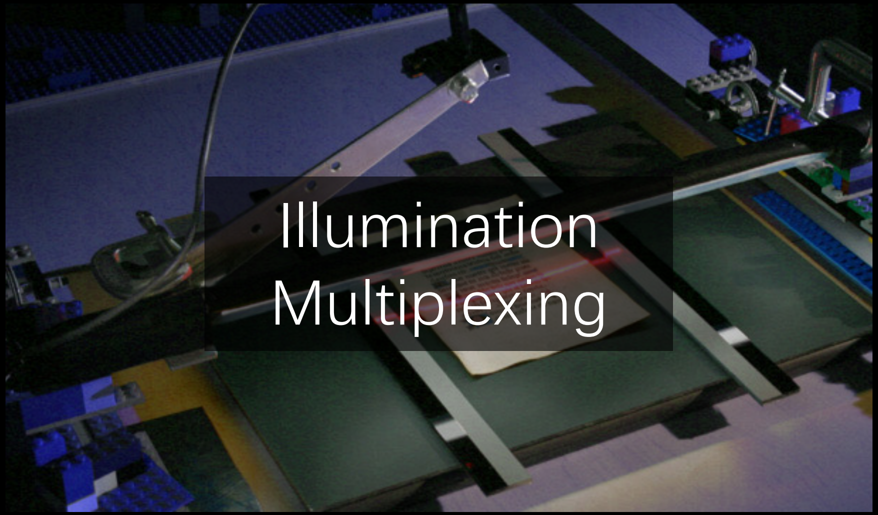
Related Work





Fixed View(s)

Unstructured Views



Fixed View(s) - Point Light(s)

- [Dong et al. 2010; Aittala et al. 2015;2016; Li et al. 2017; Deschaintre et al. 2018]

- Nearly Flat Appearance
- Low Efficiency in Lighting-View Domain
  - Point Sampling

[Dong et al.2010]

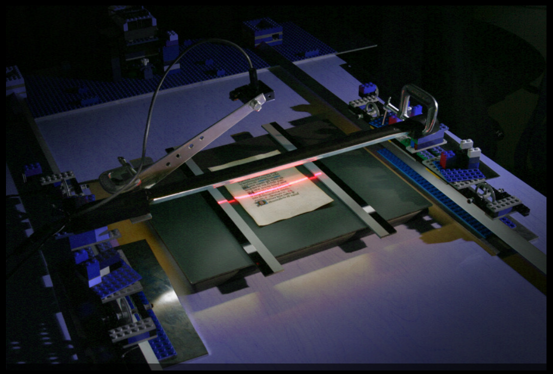
[Aittala et al.2015]

[Deschaintre et al.2018]

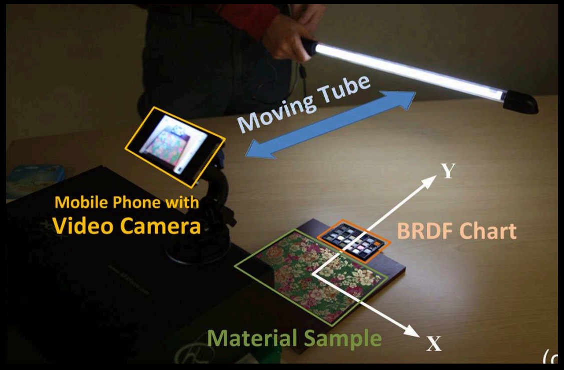
[Li et al.2017]

Fixed View(s) - Illumination Multiplexing

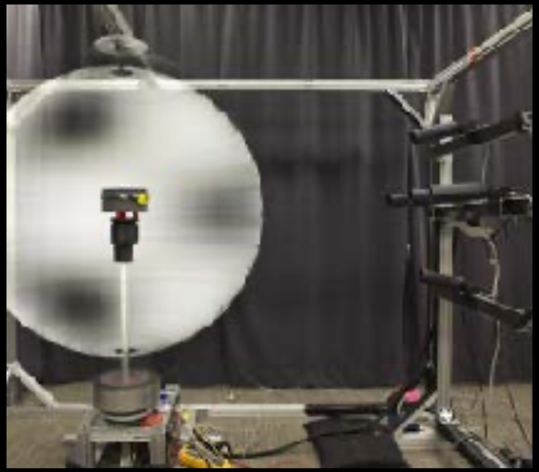
- Linear Light Source
  - [Gardner et al. 2003; Ren et al. 2011; Chen et al. 2014]
  - Planar Appearance
  - Some Requires Pre-captured BRDF Patches
- Lightstages
  - [Ghosh et al. 2009; Tunwattanapong et al. 2013; Aittala et al. 2013; Kang et al. 2019]
  - Anisotropic
- Pixel-Independent Reconstruction
- Require a Fixed View
  - No Information Aggregation Across Views



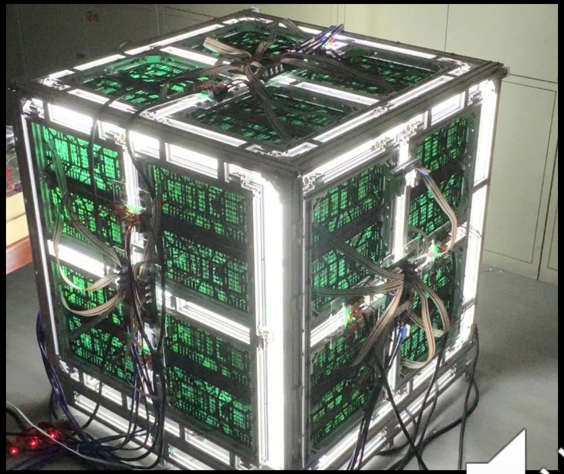
[Gardner et al. 2003]



[Ren et al. 2011]

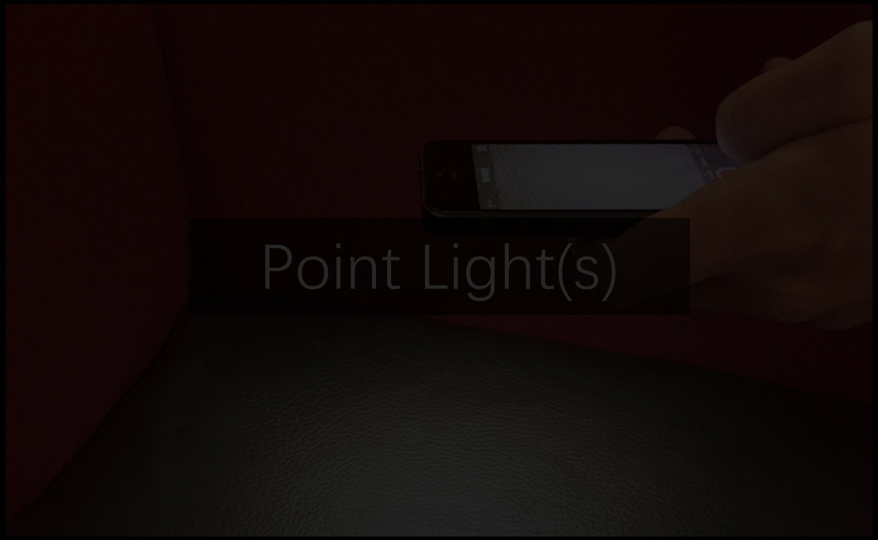


[Tunwattanapong et al. 2013]



[Kang et al. 2019]

Related Work



Fixed View(s)

Unstructured Views

Unstructured Views - Traditional Priors

- Camera-Flash [Lensch et al. 2003; Riviere et al. 2016; Nam et al. 2018]
- Kinect Sensor [Wu et al. 2015]

- Require Spatial Coherence for Regularization
  - e.g. Linear Combinations of Basis Materials
- Isotropic Reflectance

[Lensch et al. 2003]

[Riviere et al. 2016]

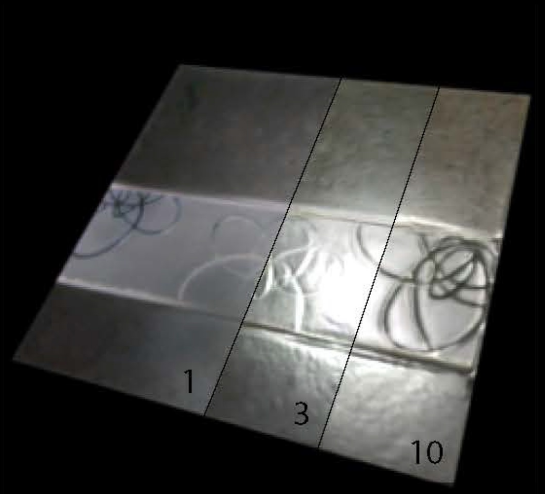
[Nam et al. 2018]

[Wu et al. 2015]

Unstructured Views - Deep-Learning-Based Priors

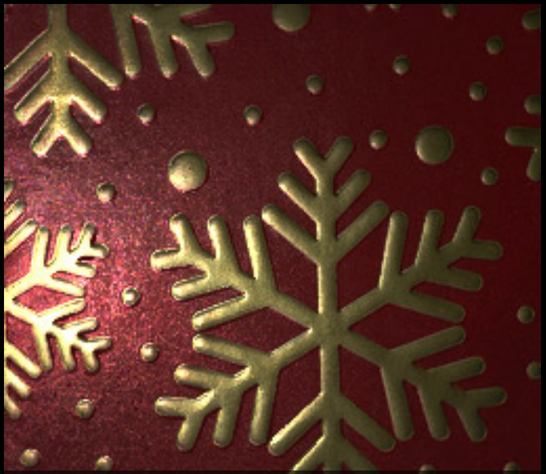
- [Deschaintre et al.2019; Gao et al.2019; Guo et al.2020; Bi et al. 2020]

- Unclear How to Extend to Complex Lights
- Often Discard View Conditions
- Isotropic Reflectance



[Deschaintre et al.2019]


[Gao et al.2019]


[Guo et al.2020]

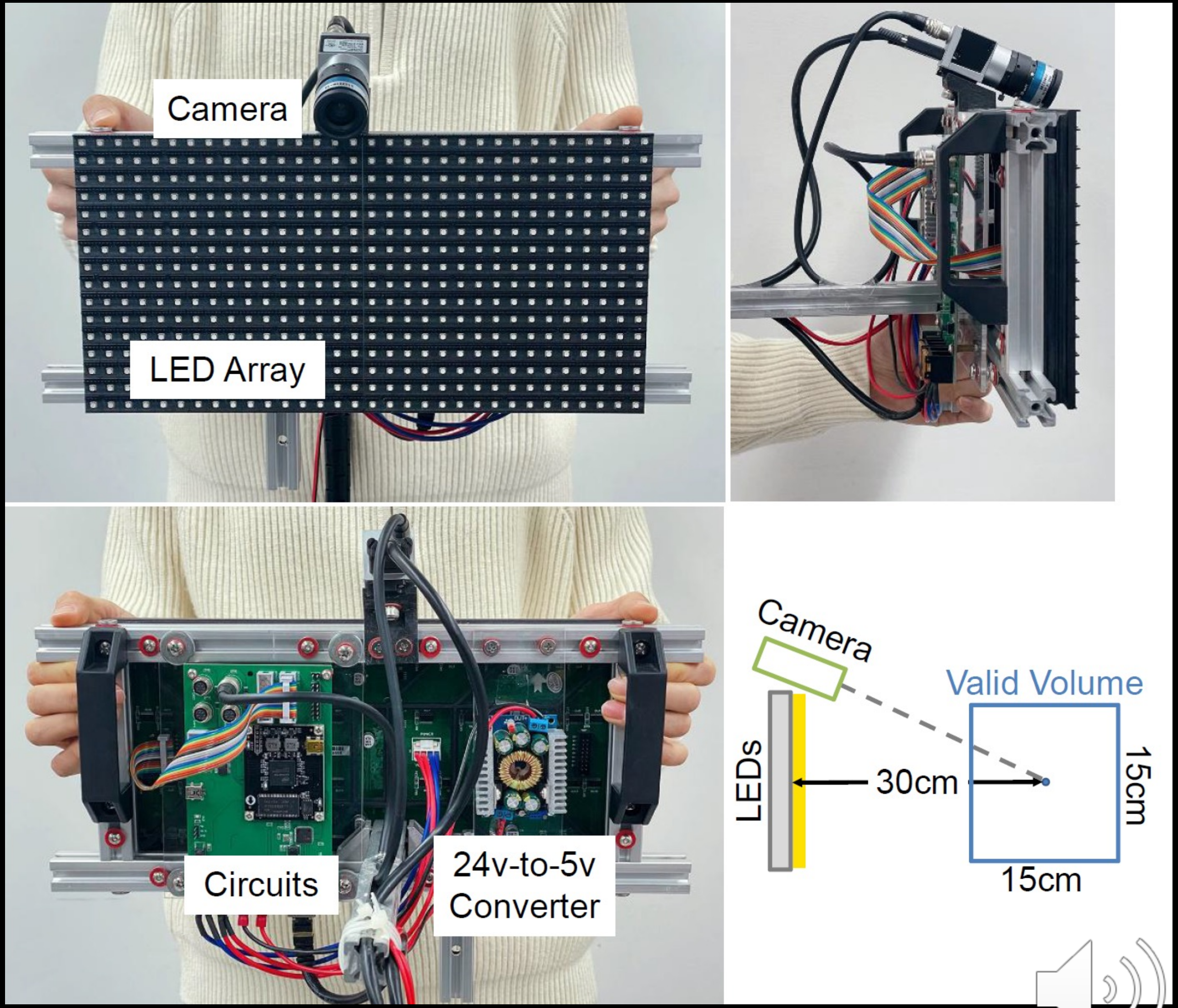

[Bi et al.2020]

SIGGRAPH 2021
VIRTUAL 9-13 AUGUST

Prototype Scanner

Prototype Scanner

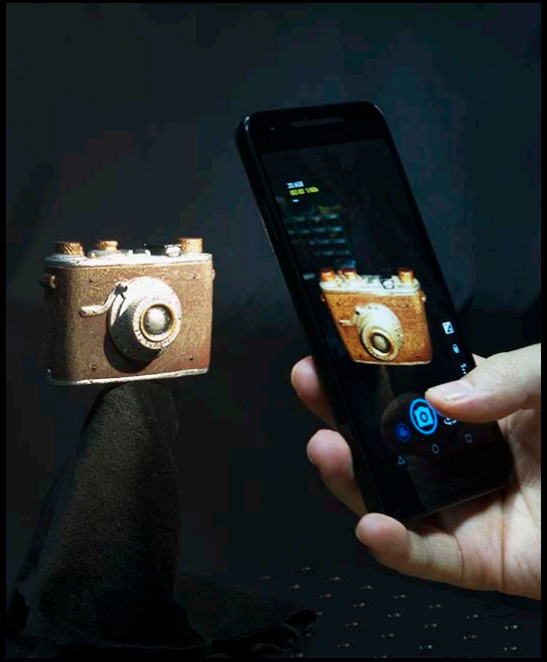
- LED Array
  - 512 Lights
  - 32cm×16cm
  - 40W
- Single Camera
  - Basler acA2440-75uc
  - 75fps
  - Resolution 2448×2048
- High-Precision Synchronization
  - Custom-designed Circuits/FPGA



Prototype Scanner

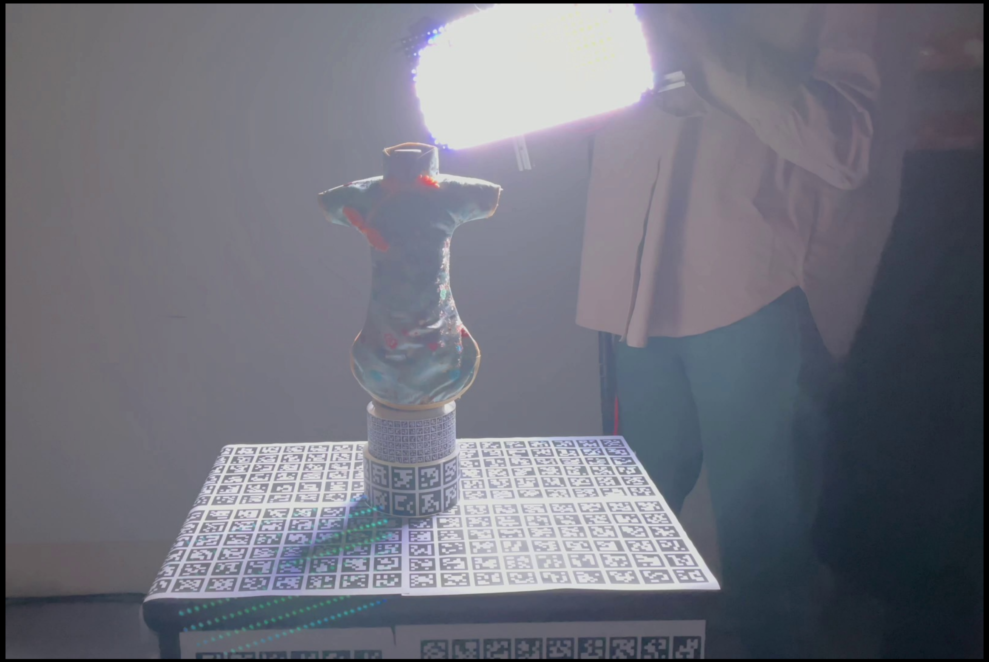
- Why an LED Array ?

Point Light



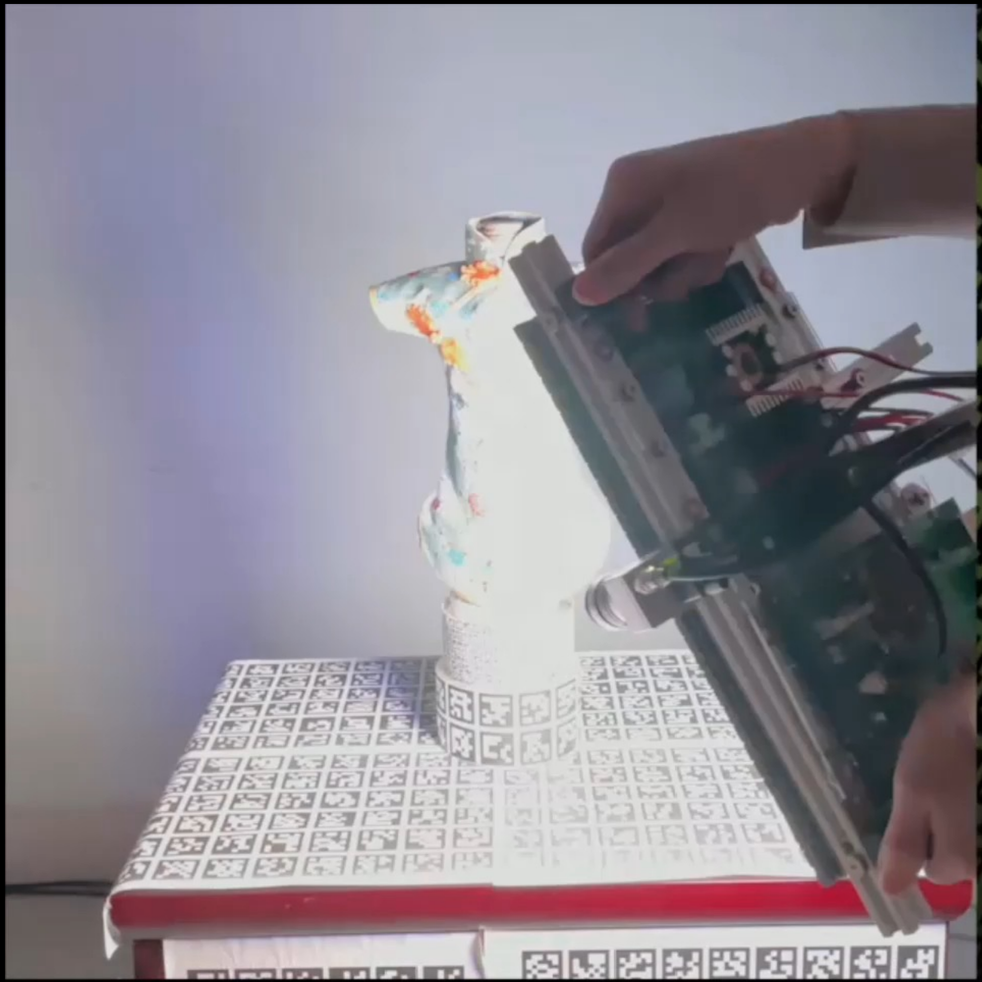
Point Sampling in Illumination Domain

LED Array



Sample Multiple Lights Simultaneously

Prototype Scanner



Appearance Acquisition Scene

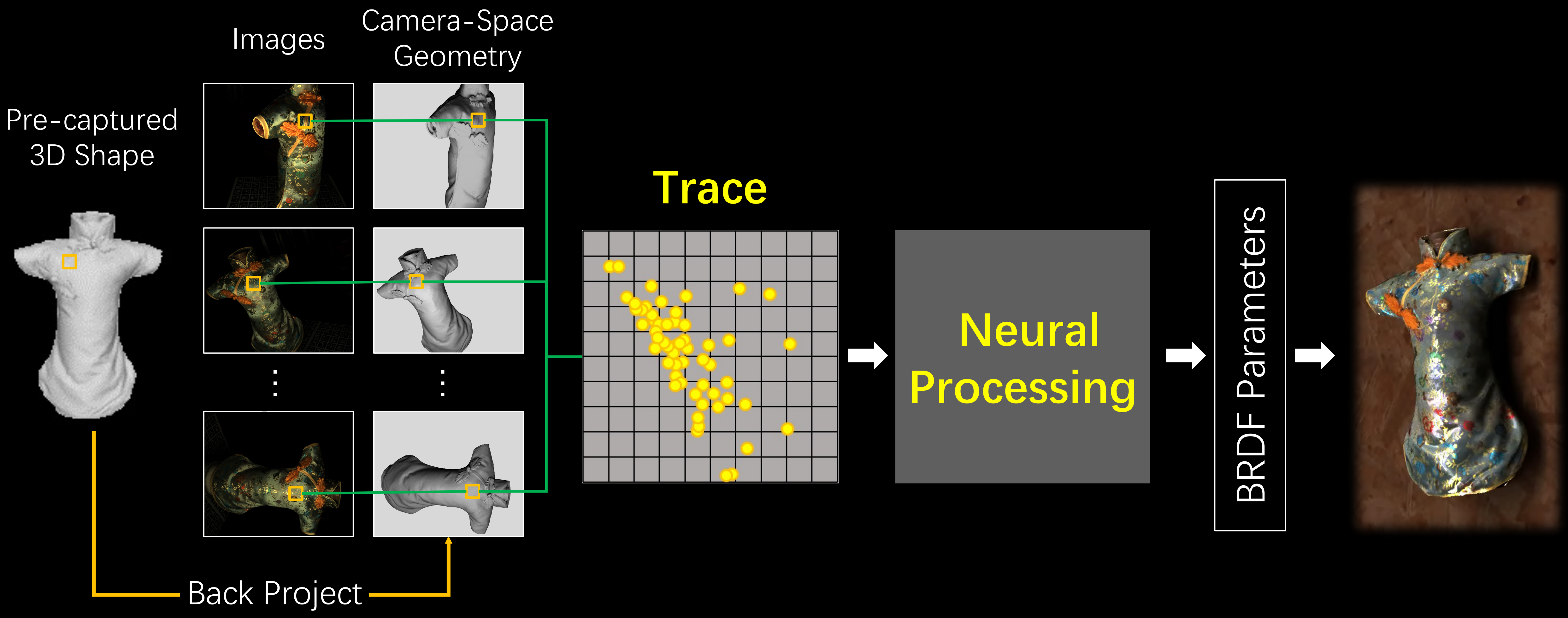
Captured Images

Our Framework

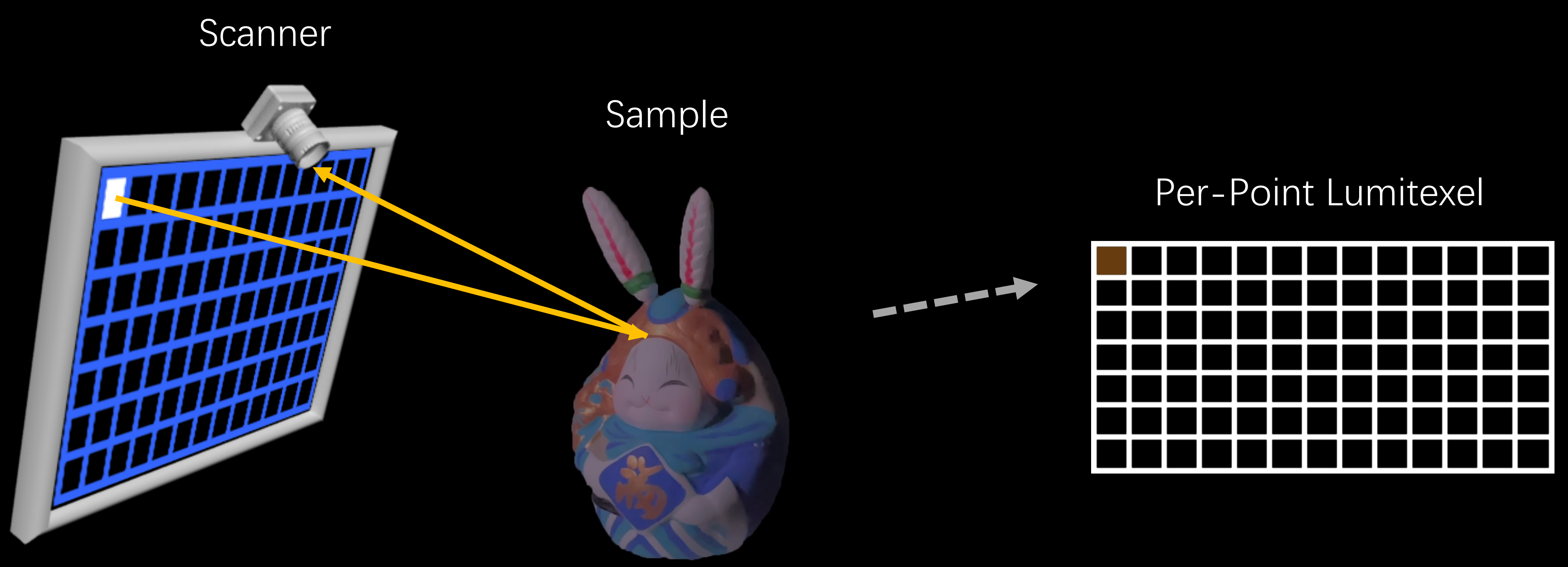
Assumptions

- Pre-Captured 3D Shape
- **Pixel-Independent** Reconstruction
- Fixed Lighting Pattern
- Relative Motion
  - Fixed Scanner / Moving Sample

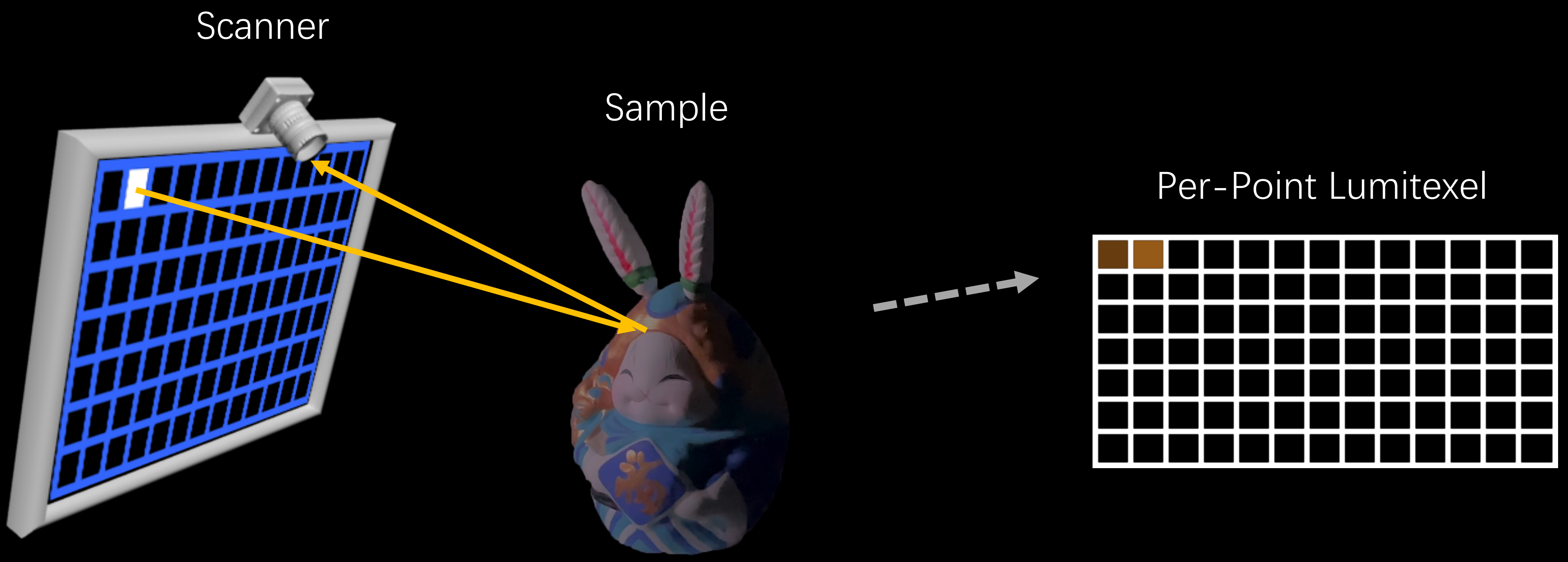
Neural Trace Photography



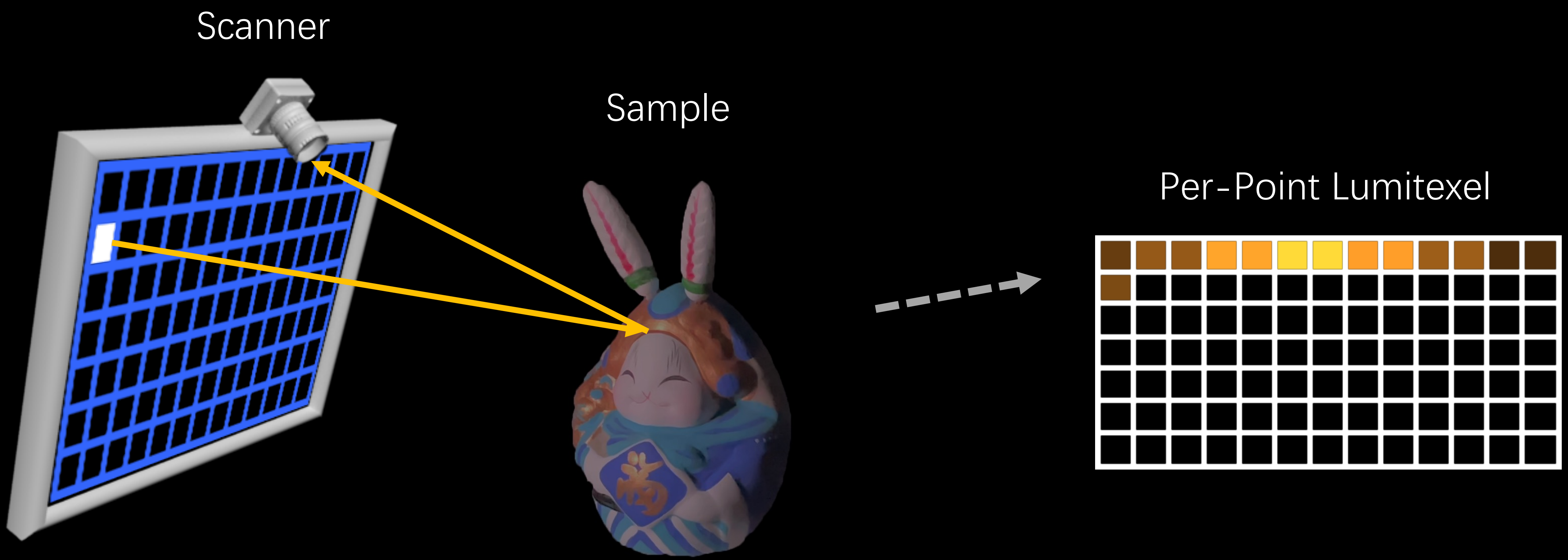
Lumitexel



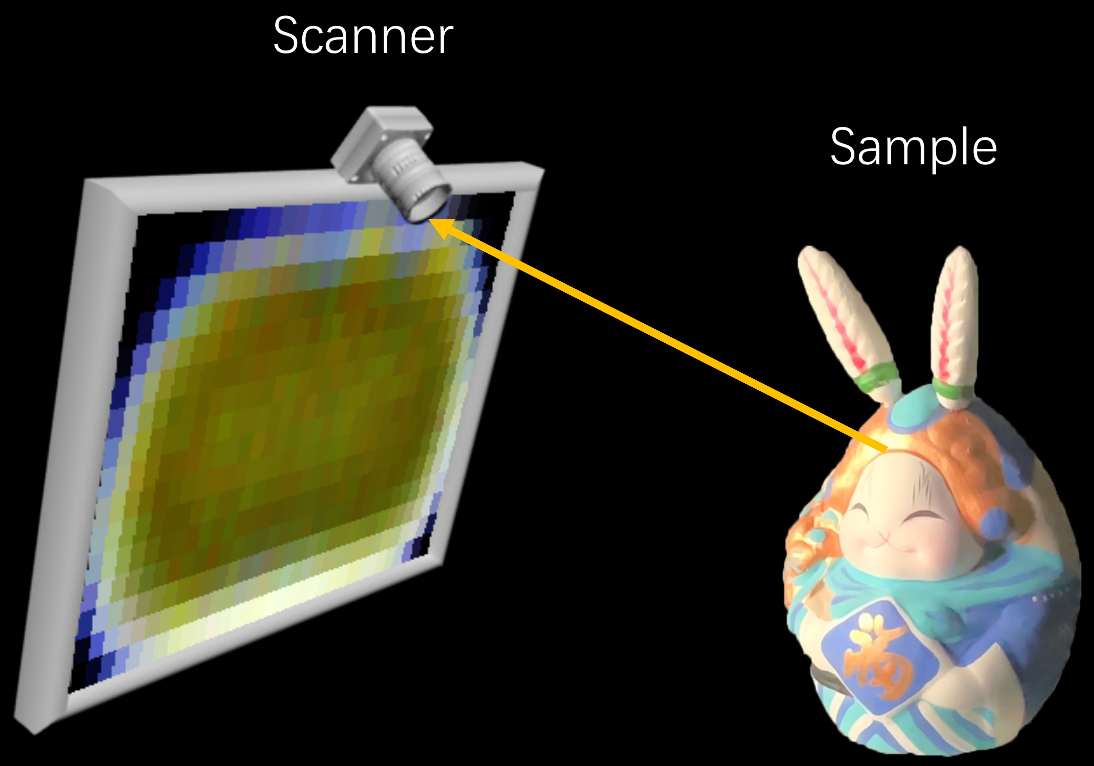
Lumitexel

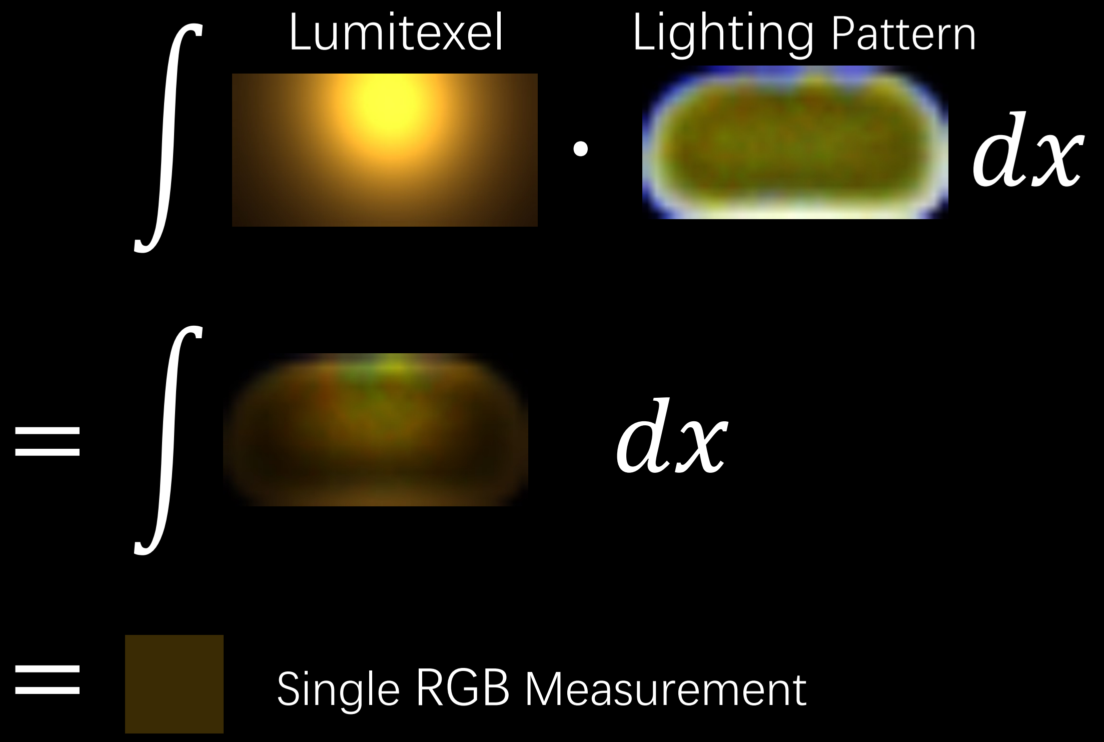


Lumitexel



Illumination Multiplexing



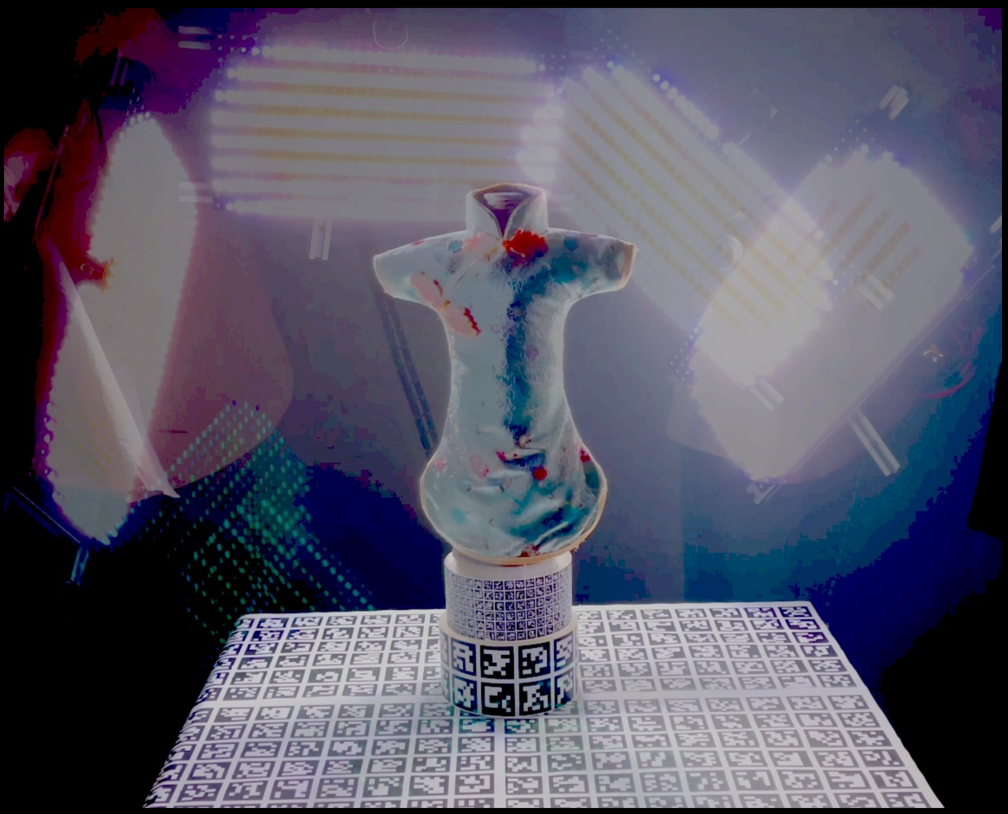


Per-point Pipeline

Per-point Pipeline

Input Trace → Network → Output Lumitexels → Fitting → BRDF Parameters

Trace



- Previous Work
  - [Dong et al. 2014; Gardner et al.2003; Ren et al. 2011; Morris and Kutulakos 2007]
- Our Definition:

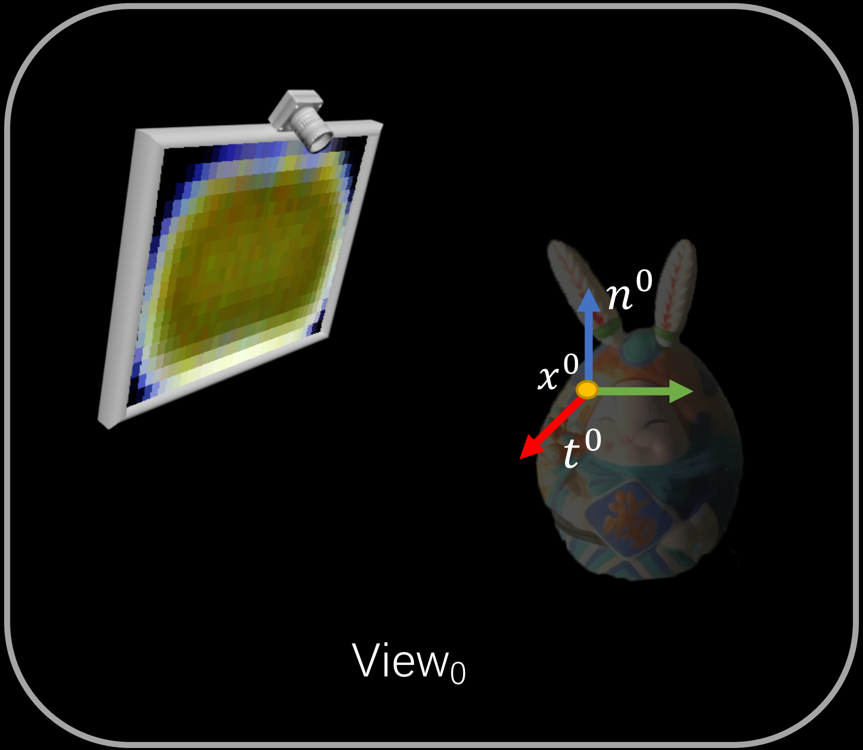
A Collection of High Dimensional Points

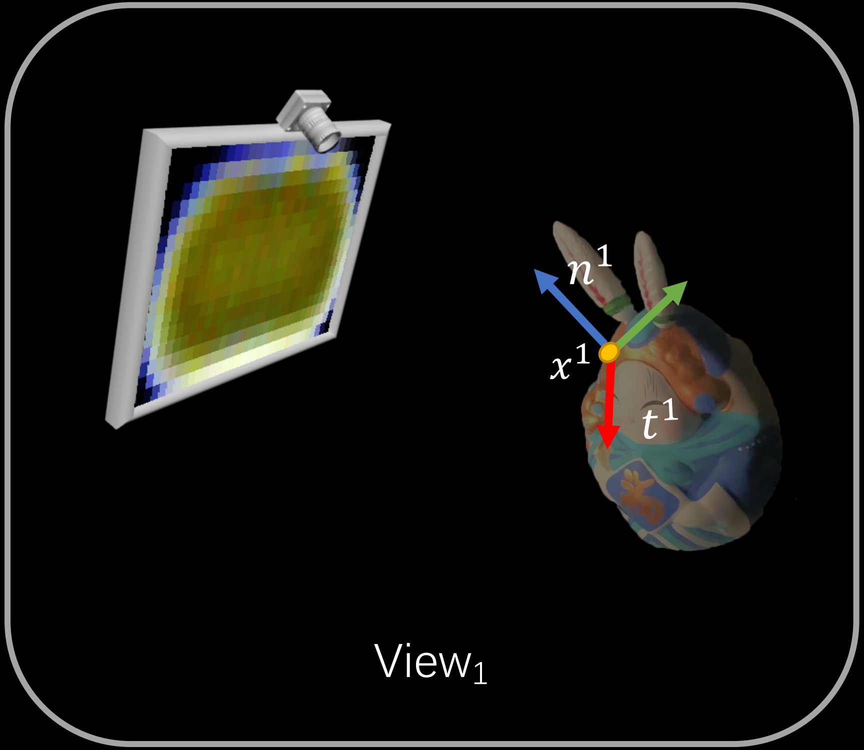
Each Point = Measure. + Acquisition Condition

Lighting Condition + View Condition

Fixed **Varying**

View Conditions





...

Trace

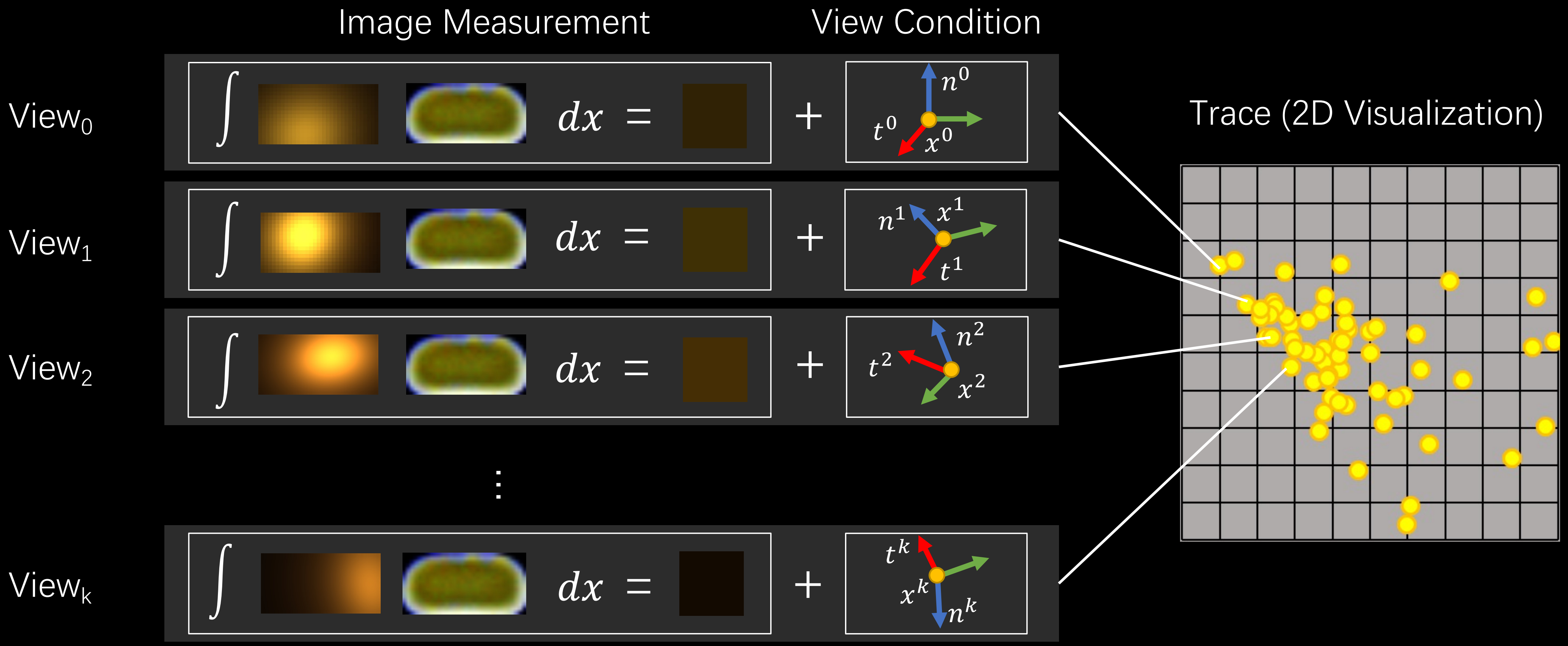
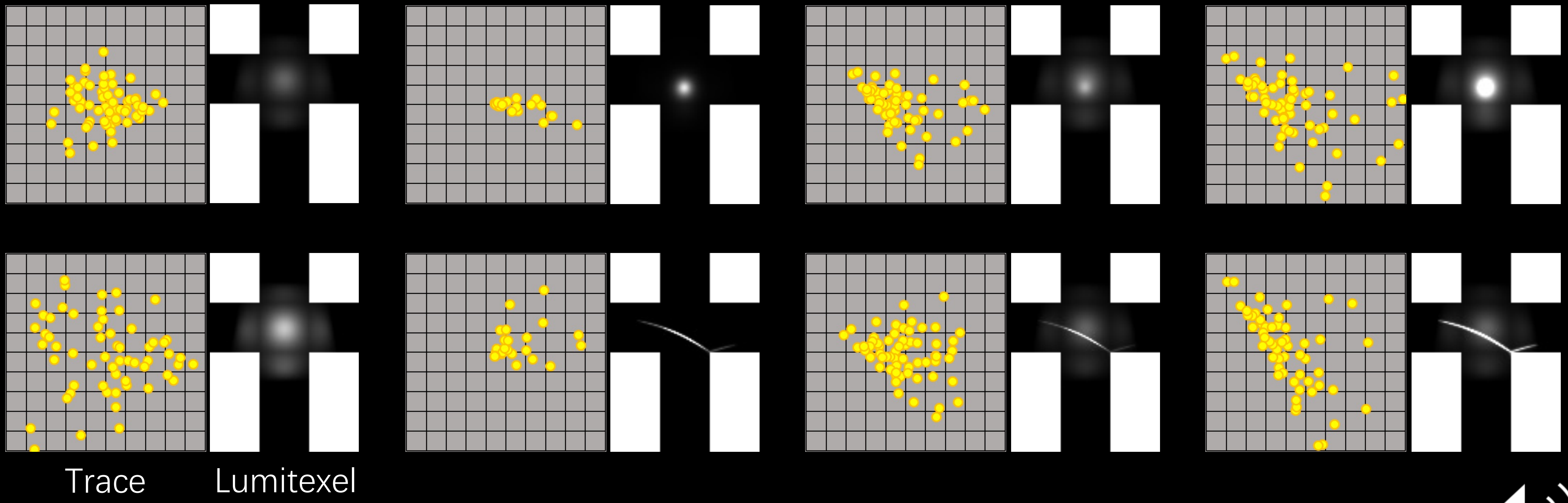


Image Measurement
View Condition
View_0
View_1
View_2
View_k
$\int$
$dx =$
+
n^0
t^0
x^0
n^1
x^1
t^1
n^2
t^2
x^2
t^k
x^k
n^k
⋮
Trace (2D Visualization)

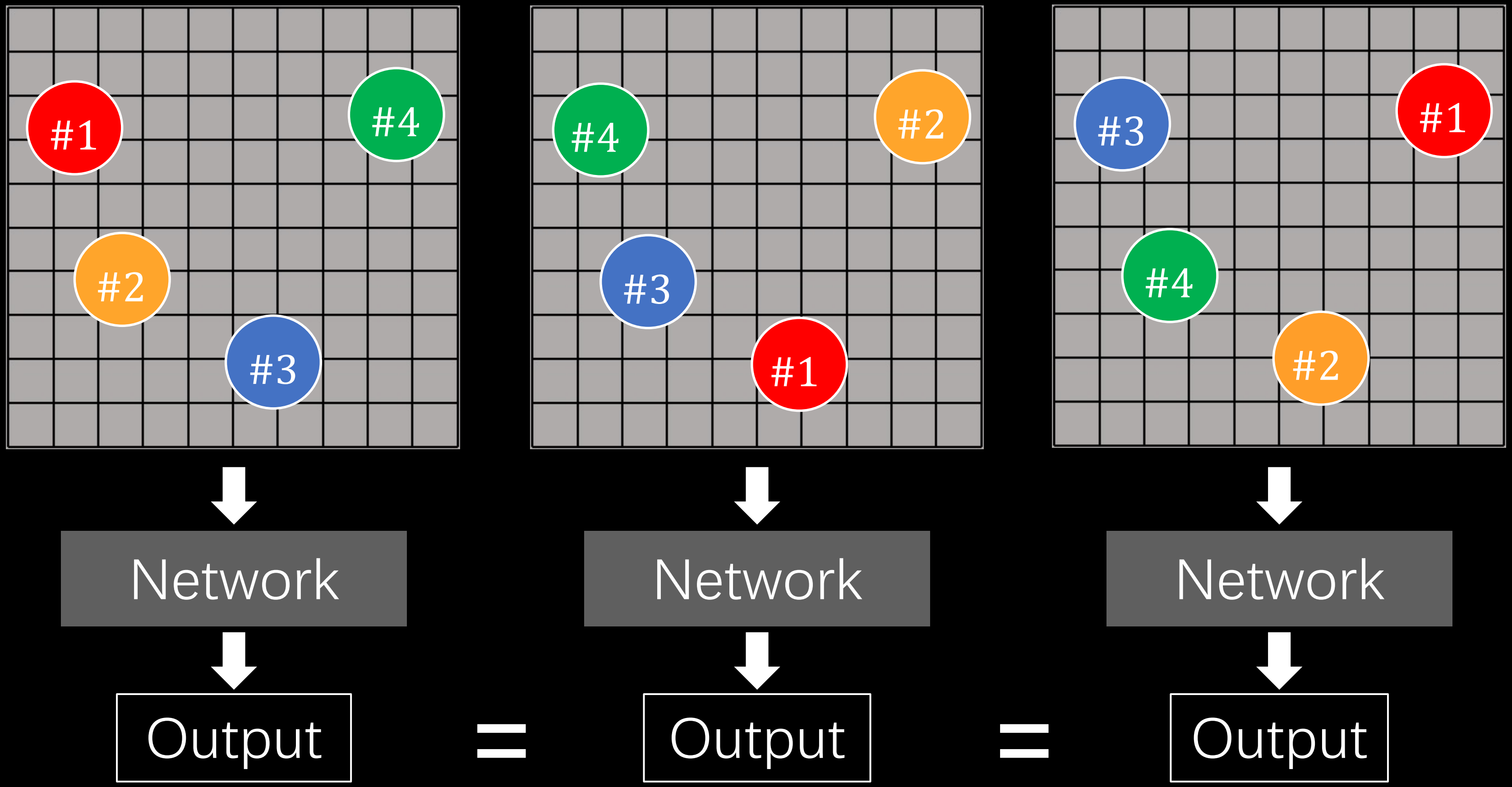
Trace

- Correlation Between Trace & Lumitexel
  - Challenging to Derive Manually



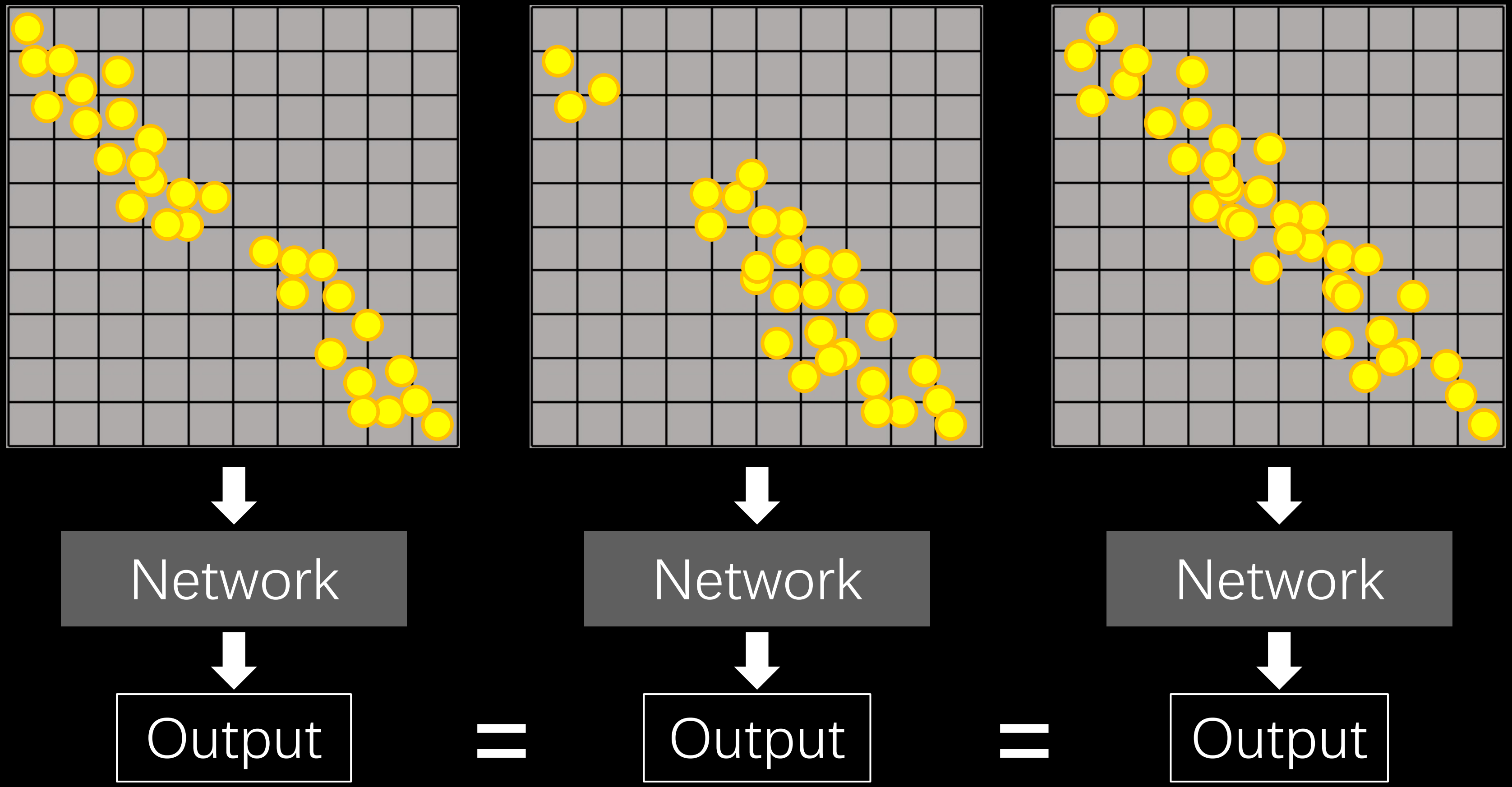
Trace

- Order Independence



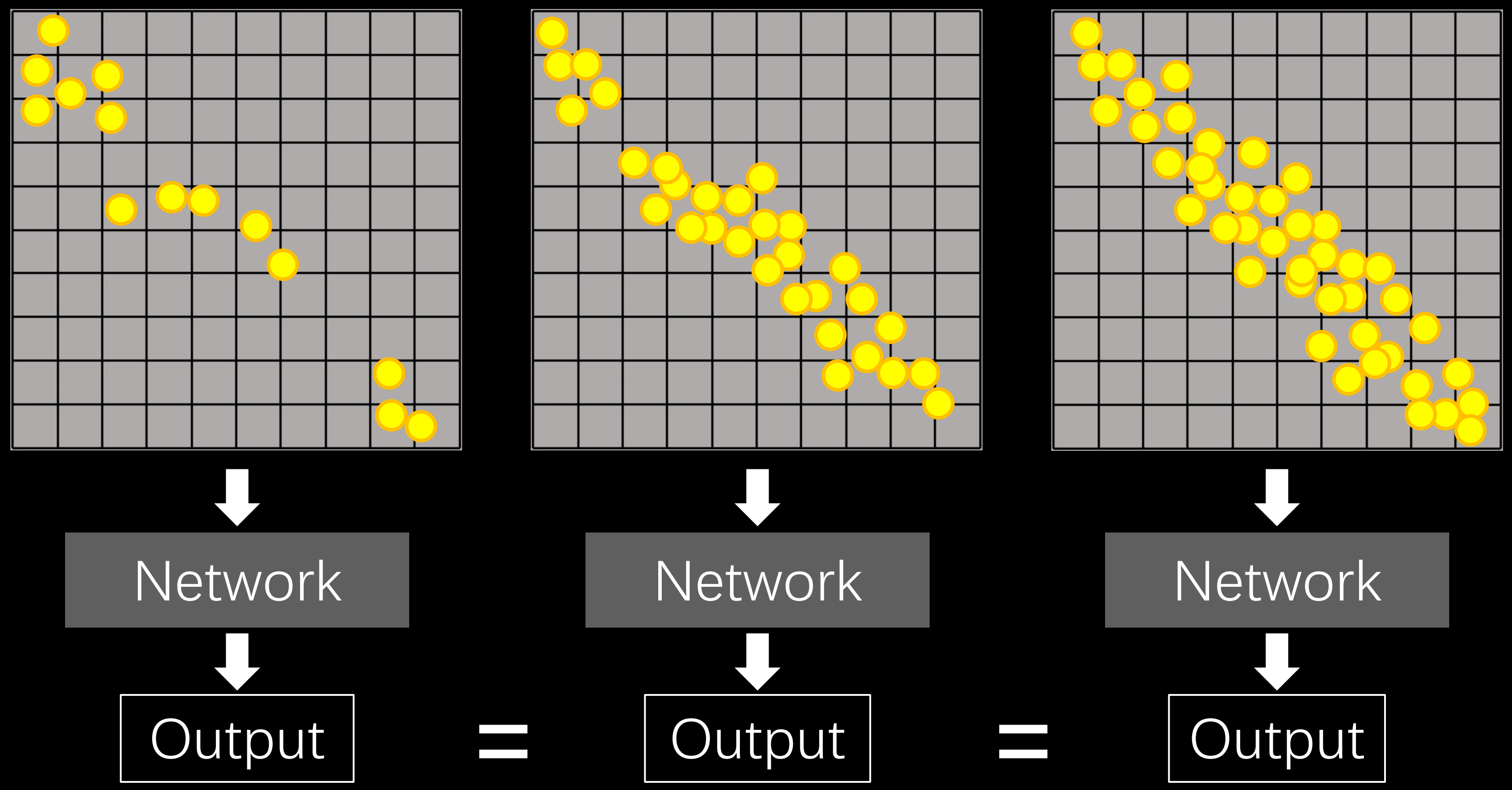
Trace

- Irregularly Sampled



Trace

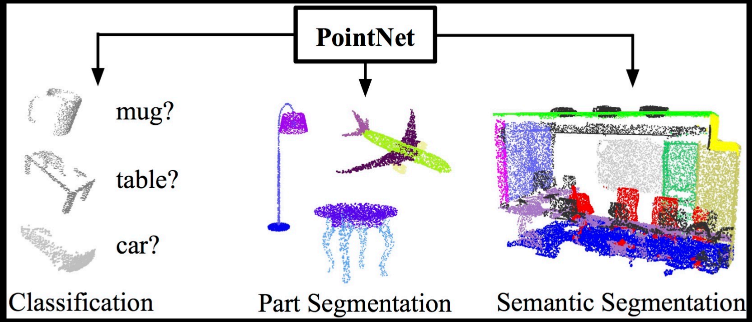
- Variable-Length



Trace

- Correlation Between Trace & Lumitexel
- Order Independence
- Irregularly Sampled
- Variable-Length

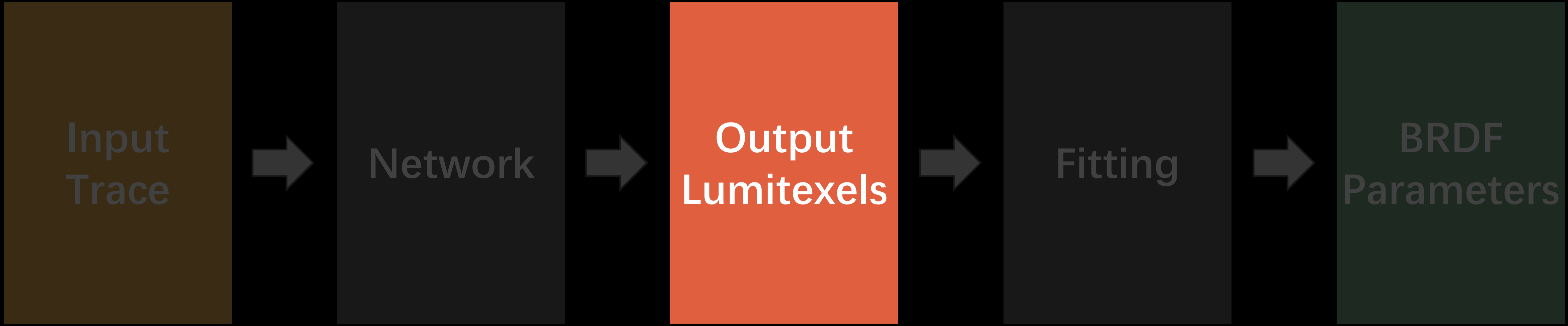
Motivate the Use of Geometry Learning Tools



[Qi et al. 2017]

Key Insight: Appearance Scanning = Geometry Learning

Per-point Pipeline



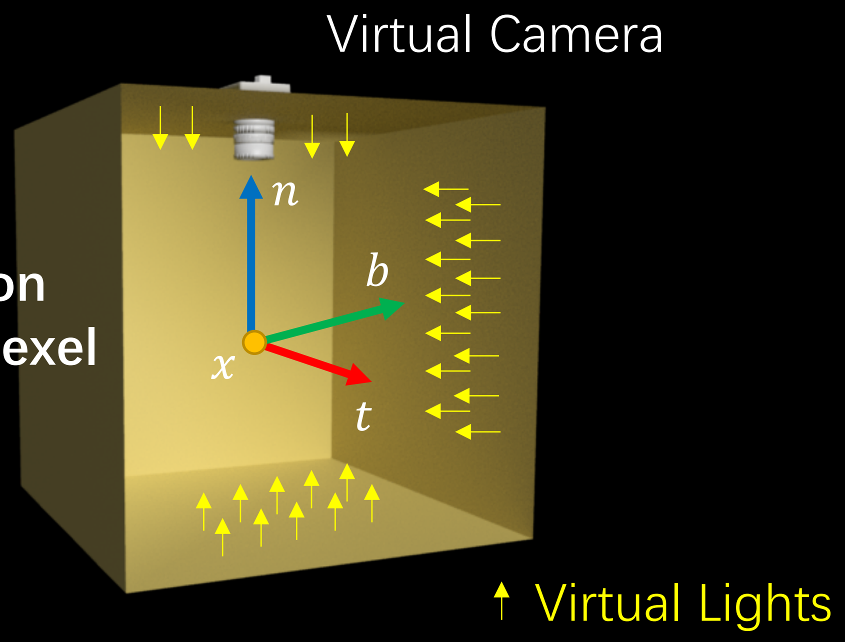
Output Lumitexel

- Challenges
  - LED Array Coverage is Incomplete
  - Multiple Unstructured Views
    - Which View Should Be the Output One?

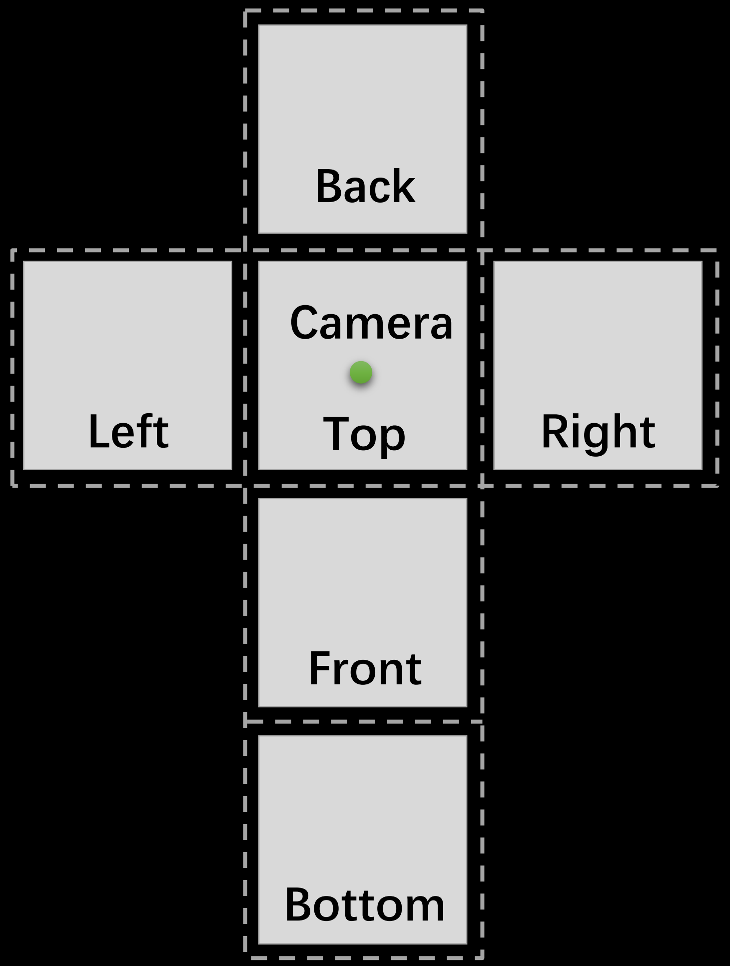
Output Lumitexel

- Virtual Camera
- Virtual Lights
  - 6×8^2 Diffuse Lumitexel
  - 6×32^2 Specular Lumitexel
- Use the Shading Frame as the Coordinate System

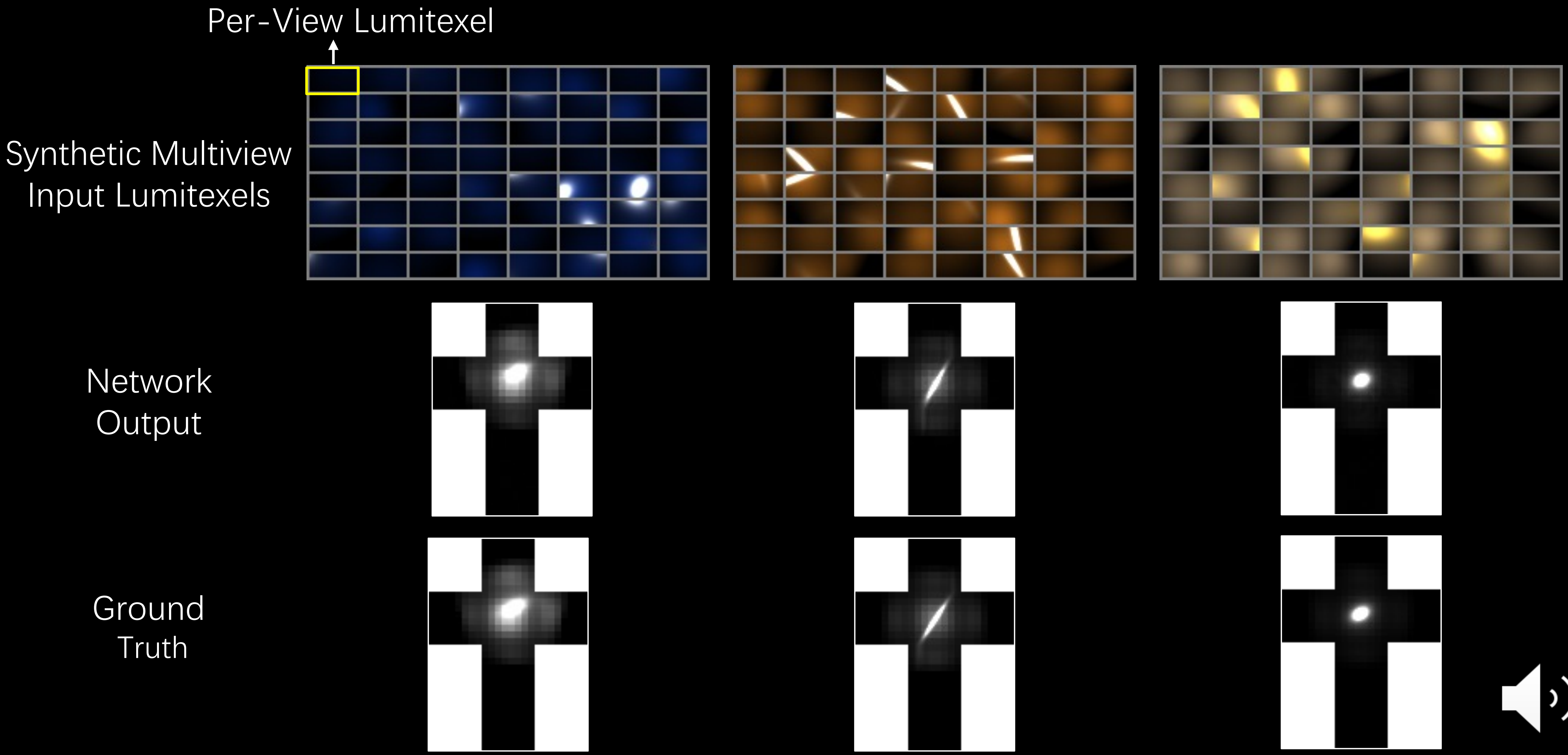
Parameterization For Output Lumitexel



Output Lumitexel Layout



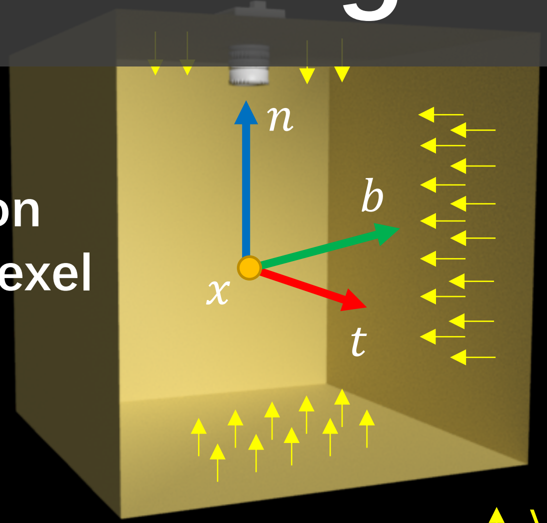
Synthetic Lumitexel Reconstruction



Output Lumitexel

- Virtual Camera
- Virtual Lights
  - 6×8^2 Diffuse Lumitexel
  - 6×32^2 Specular Lumitexel
- Use the Shading Frame as the Coordinate System

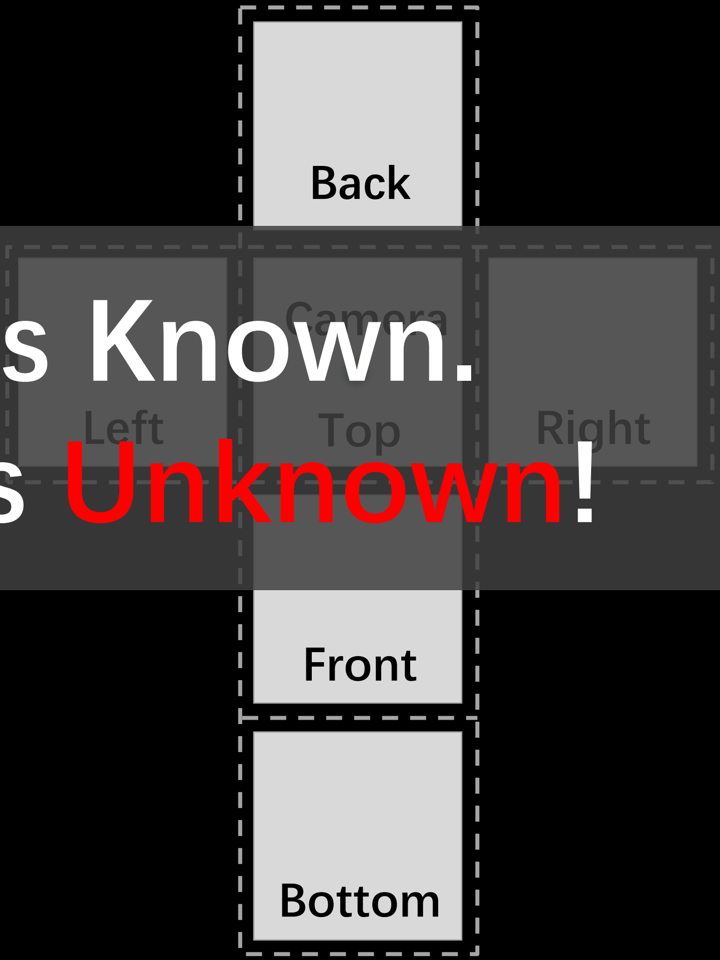
Only Geometric Frame is Known.
Accurate Shading Frame is Unknown!



Parameterization
For Output Lumitexel

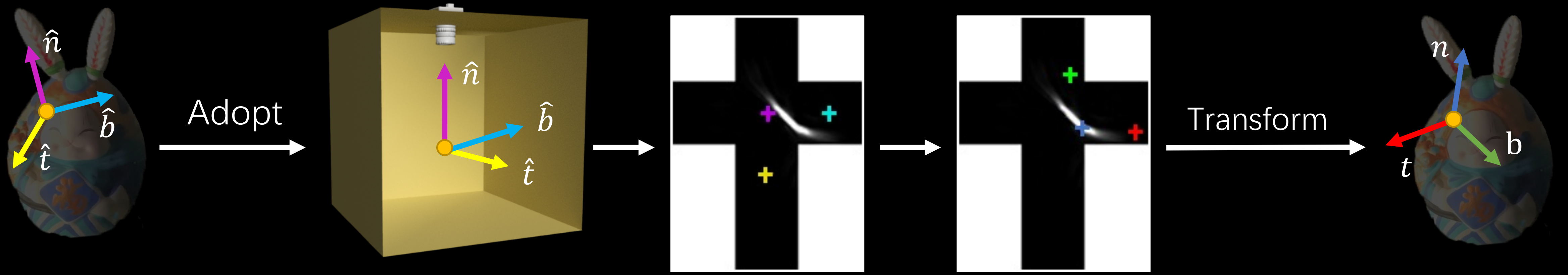
Virtual Lights

Output Lumitexel Layout

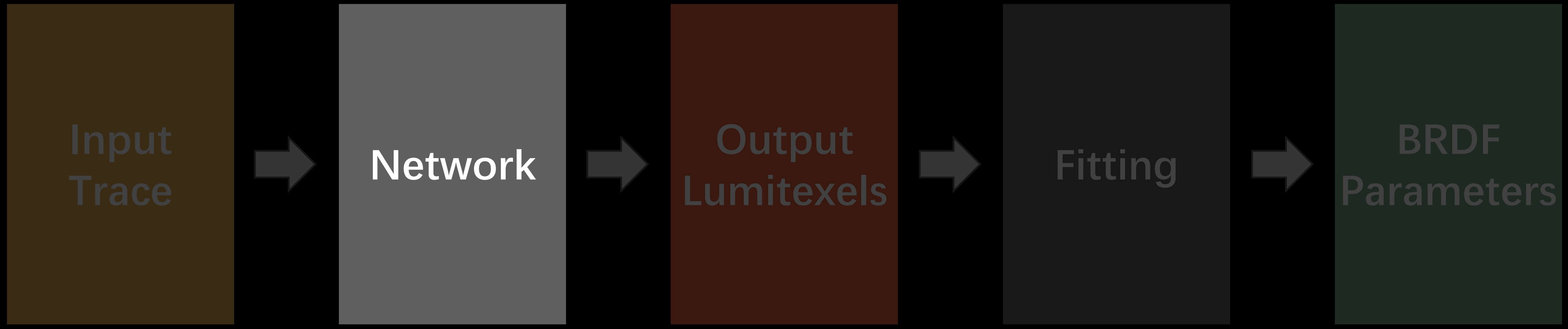


Output Lumitexel

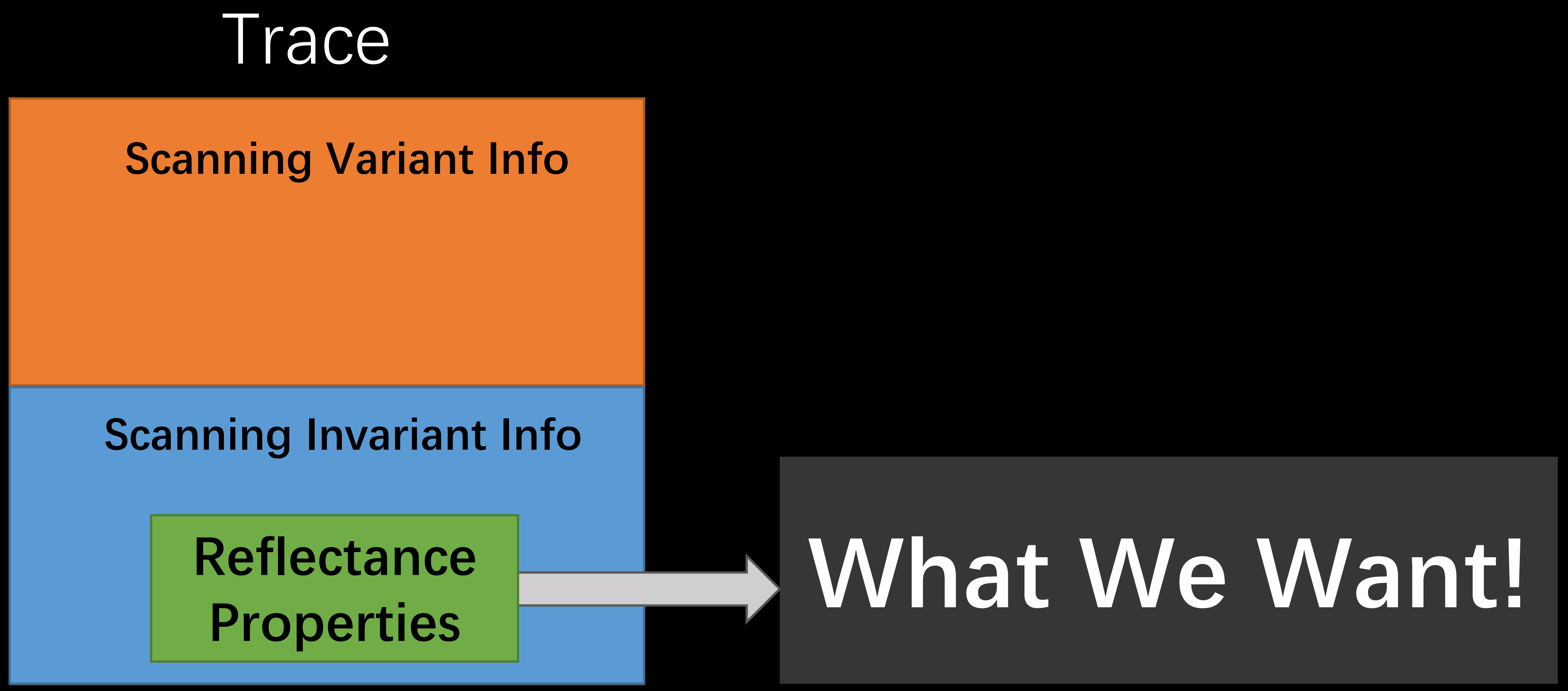
- Use Geometric Frame Instead



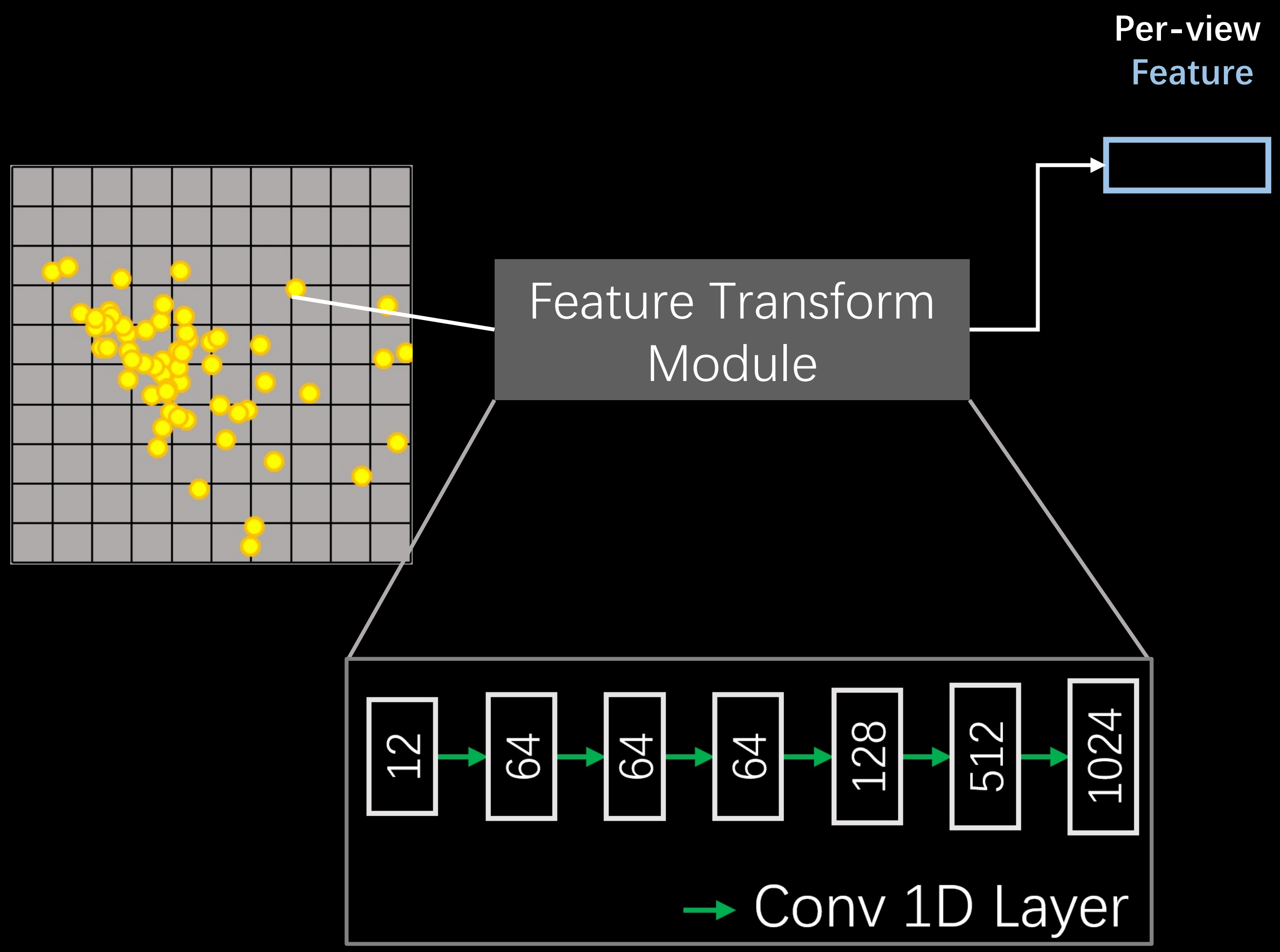
Per-point Pipeline



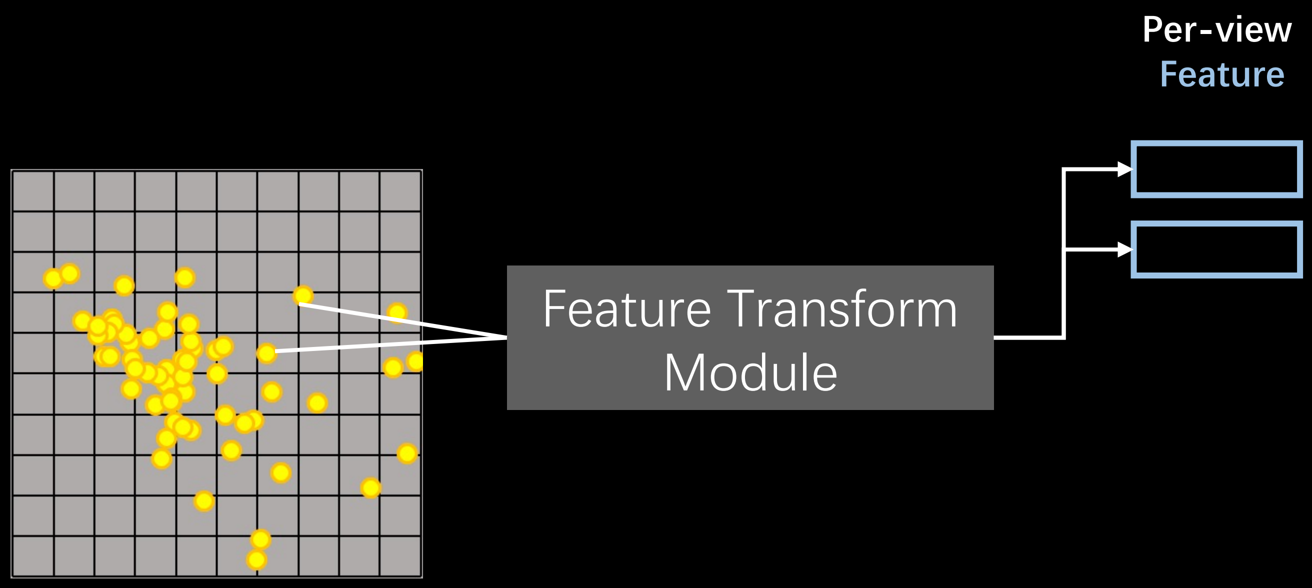
Network


Trace
Scanning Variant Info
Scanning Invariant Info
Reflectance Properties
What We Want!

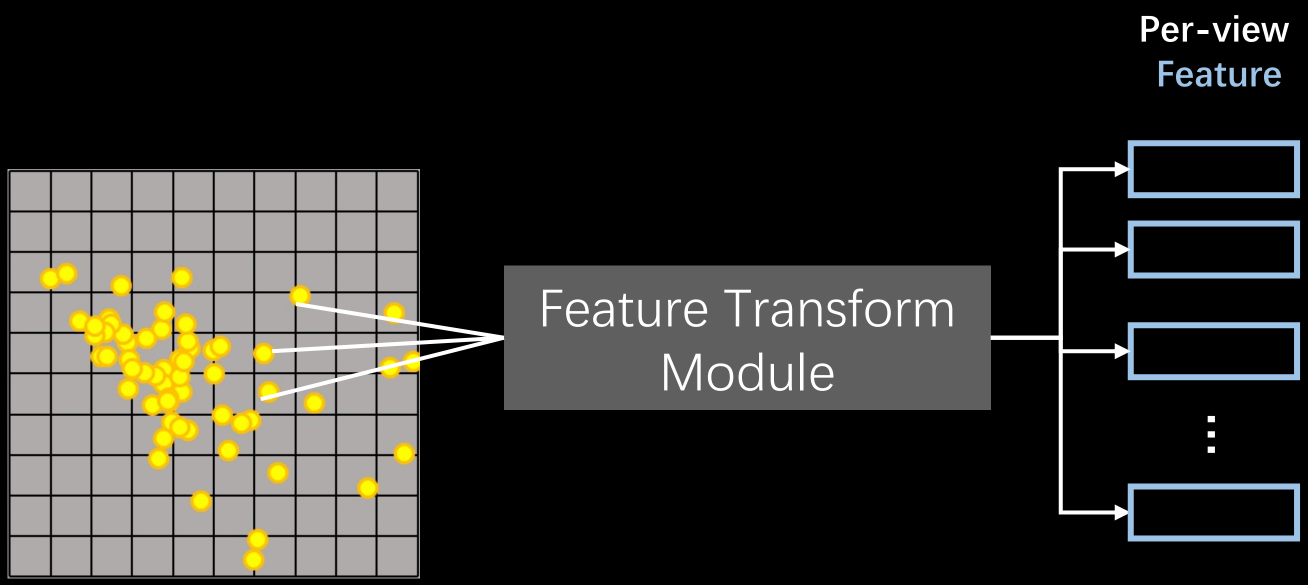
Network


Per-view
Feature
Feature Transform
Module
12
64
64
64
128
512
1024
Conv 1D Layer

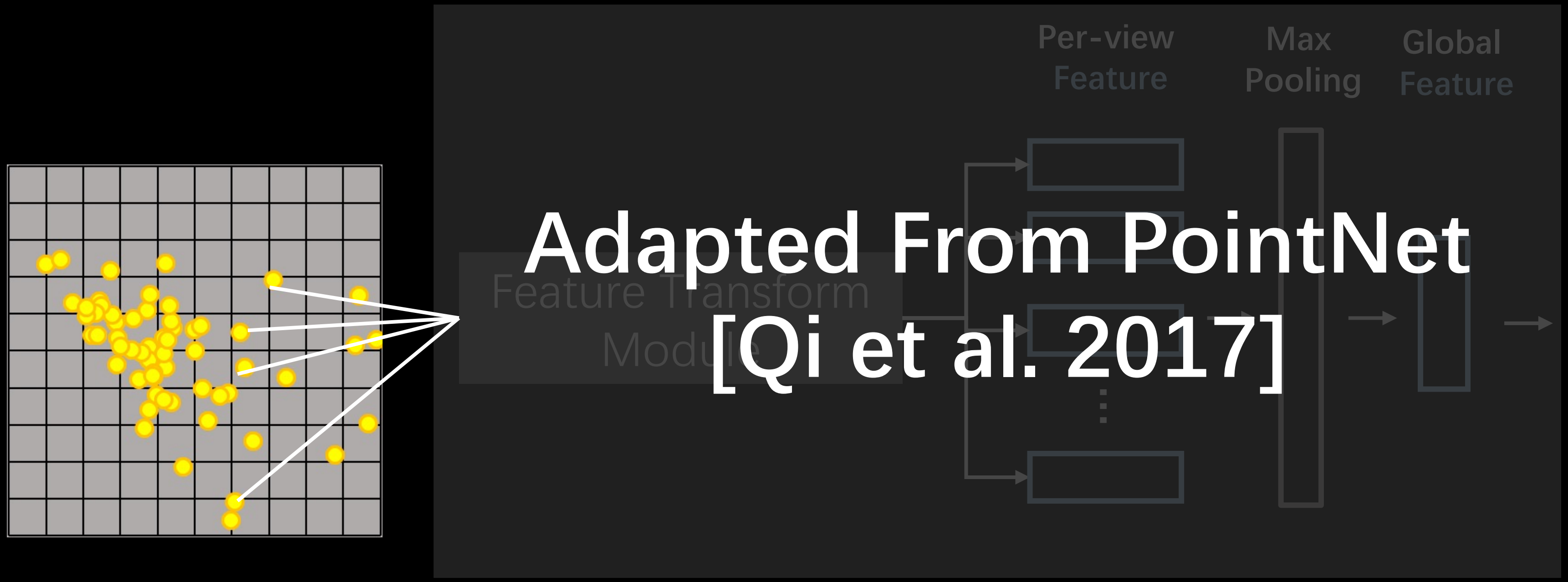
Network



Network

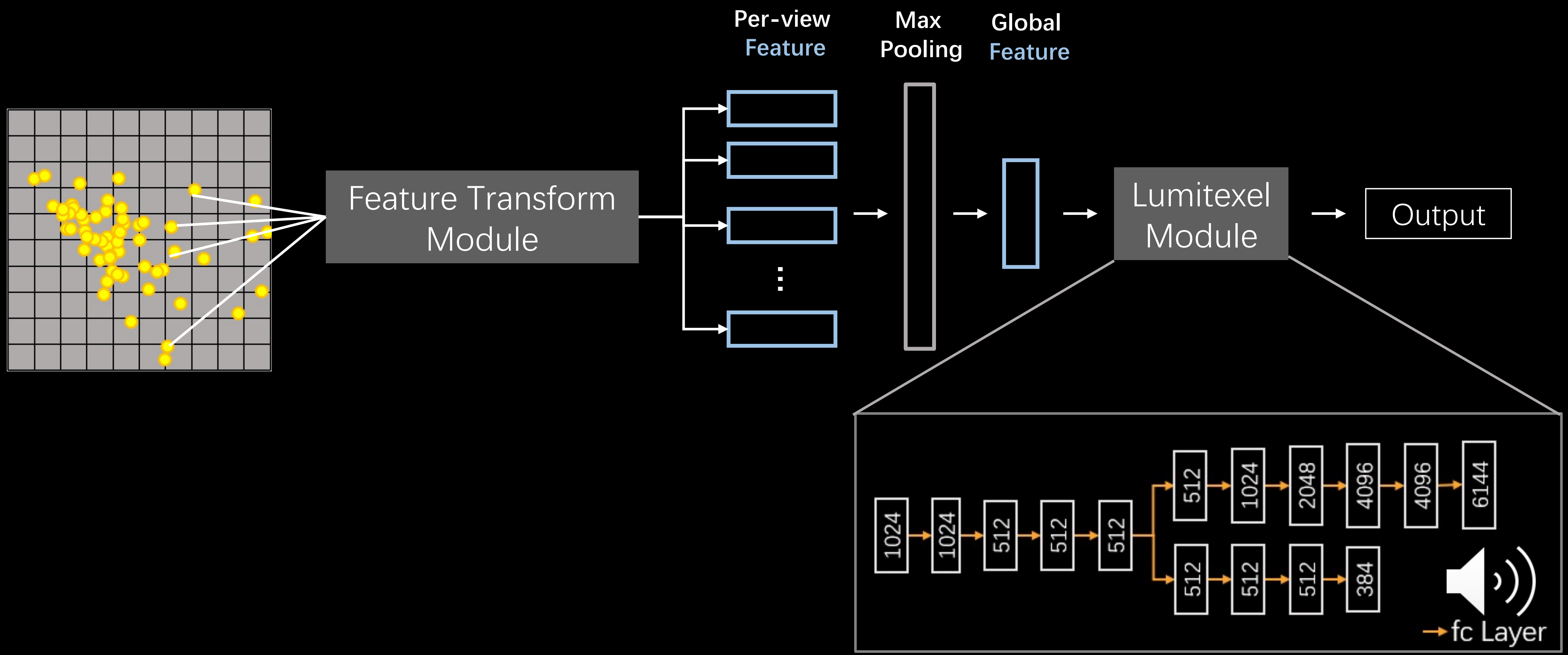


Network



Adapted From PointNet [Qi et al. 2017]

Network

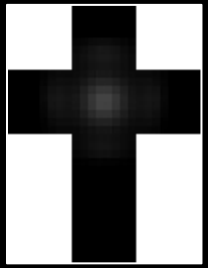
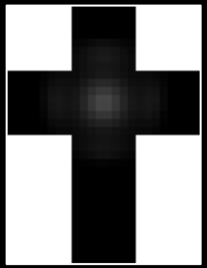


Per-view Feature
Max Pooling
Global Feature
Feature Transform Module
Lumitexel Module
Output
1024
1024
512
512
512
512
1024
2048
4096
4096
6144
512
512
512
384
fc Layer

Network

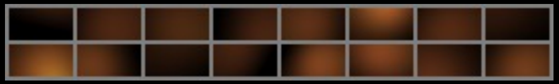
- Loss Function

$L = E_{Diffuse} + \beta \cdot E_{Specular}$

$= \lambda_d \cdot \|$  $-$  $\|^2 + \beta \cdot \lambda_s \|$  $-$  $\|^2$

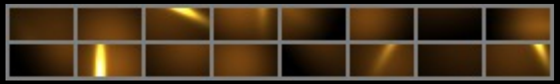
Ground Truth — Output — Ground Truth — Output

β: Confidence = Input Highlight Coverage



$\beta = 0.5$

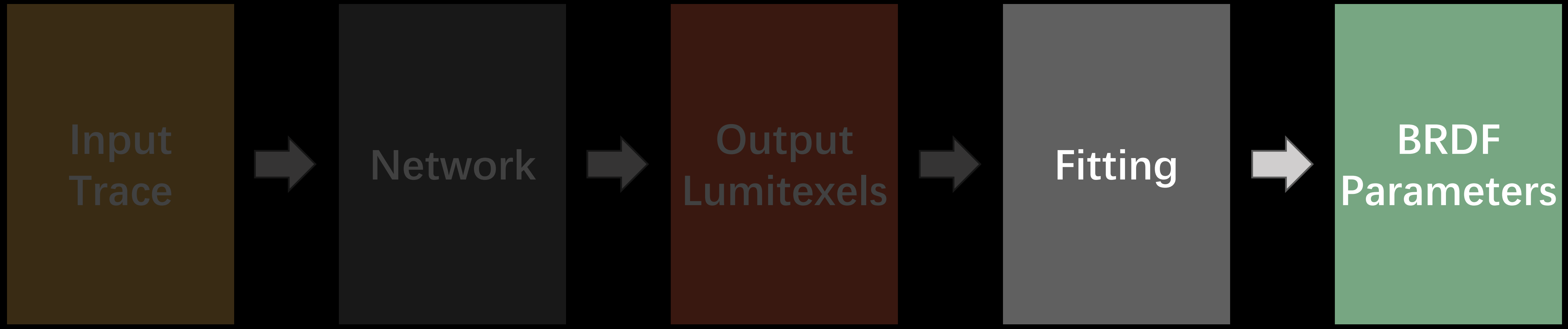
$\beta = 0.75$



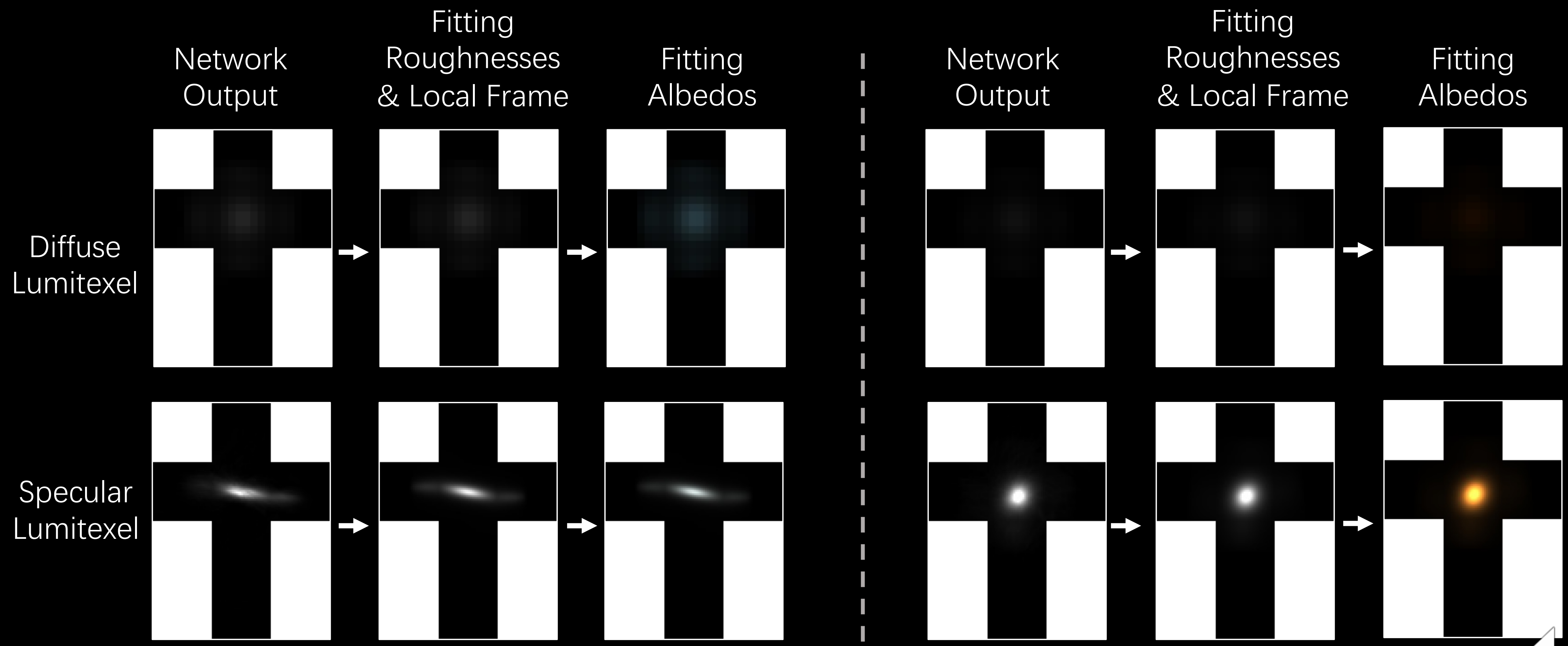
$\beta = 1.0$

Less Coverage → More Coverage

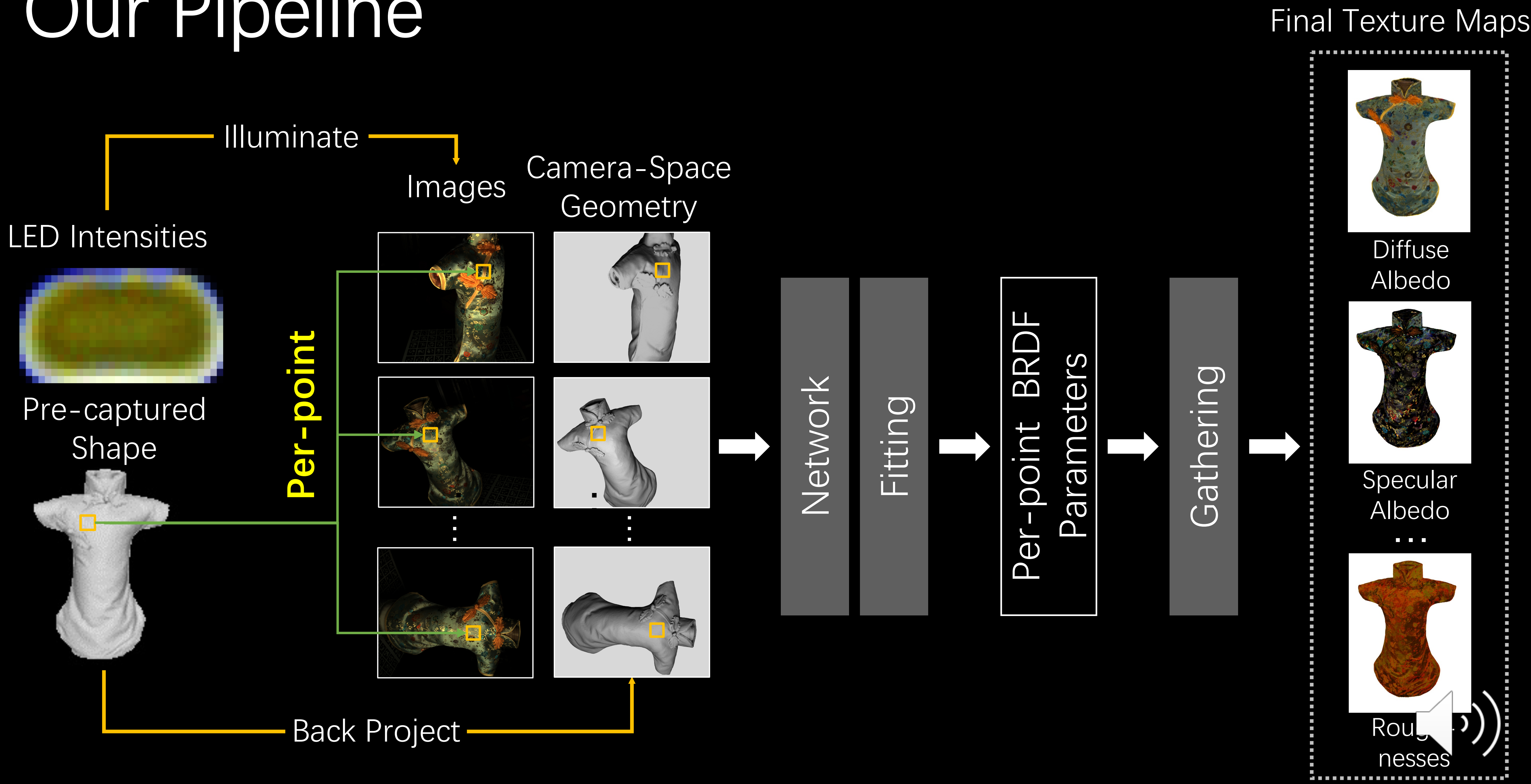
Per-point Pipeline



Fitting


Network Output
Fitting Roughnesses & Local Frame
Fitting Albedos
Network Output
Fitting Roughnesses & Local Frame
Fitting Albedos
Diffuse Lumitexel
Specular Lumitexel

Our Pipeline



Training Data

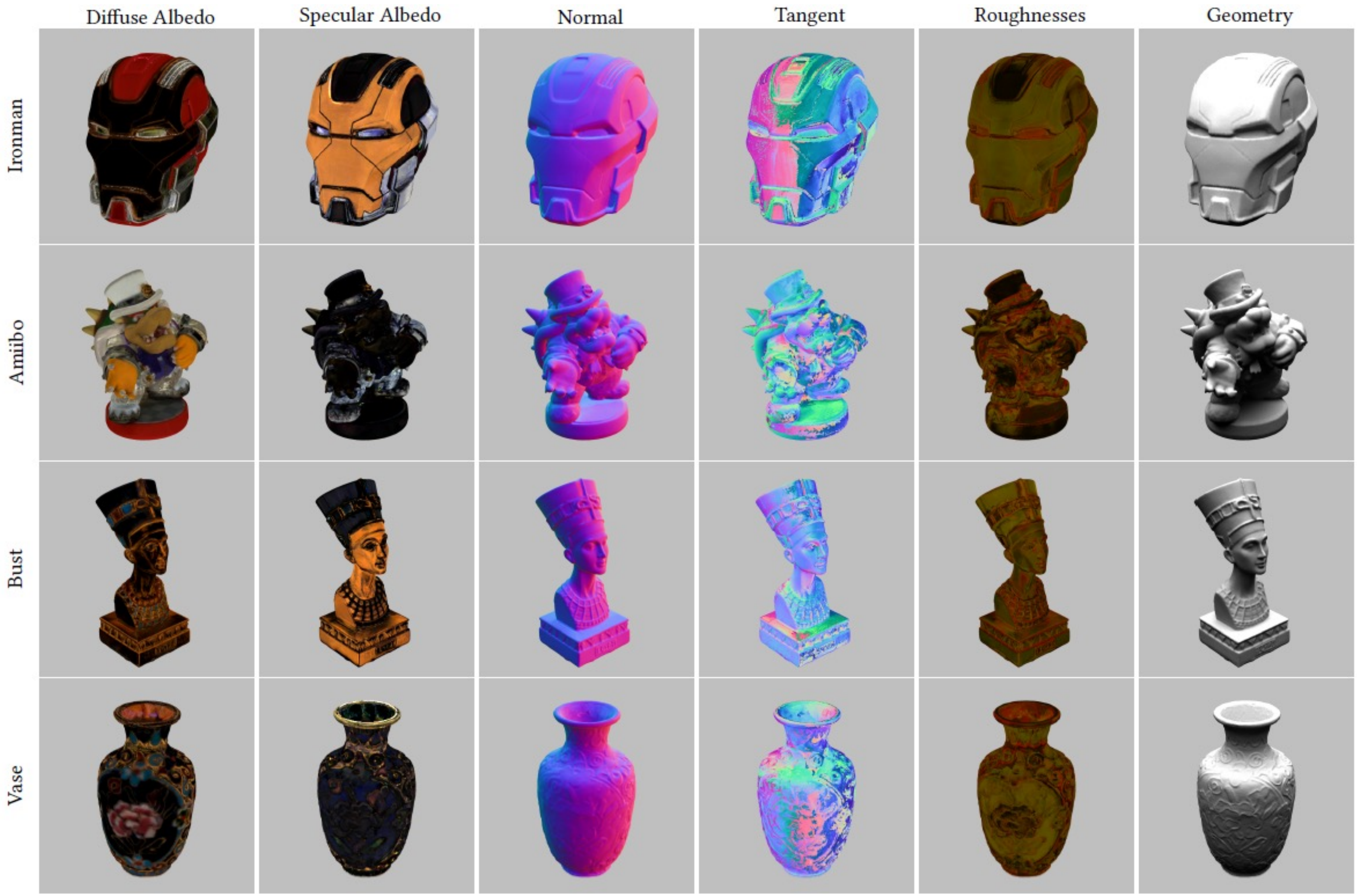
- 200M Synthetic Traces
  - Random BRDF Parameters (Anisotropic GGX)
  - Random Position / Visible Local Frame for Each View
- To Increase Robustness
  - Add Gaussian Noise to BRDF Parameters / Simulated Measurements
  - 30% Dropout Rate to fc Layers

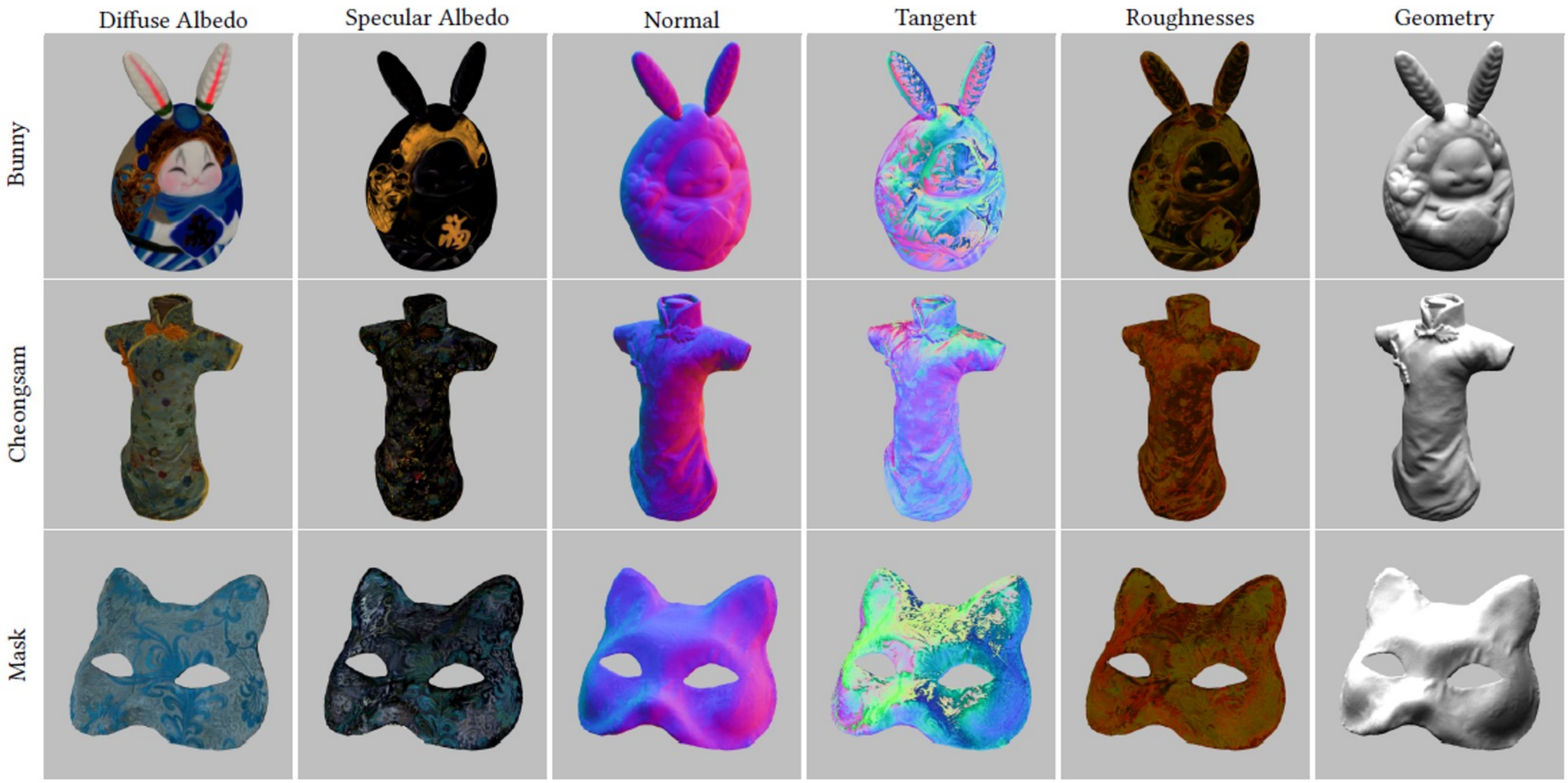
Statistics

Max Dimension of a Sample	9~32	cm
Shape Scanning	20	minutes
Appearance Scanning	9	minutes (1,000 photos)
Image Registration	2	hours
Lumitexel Prediction	6	minutes
Reflectance Fitting	2	hours
Training	66	hours

Results

Captured Appearance Rendered with Novel Lighting & View Conditions


Diffuse Albedo
Specular Albedo
Normal
Tangent
Roughnesses
Geometry
Ironman
Amiibo
Bust
Vase


Diffuse Albedo
Specular Albedo
Normal
Tangent
Roughnesses
Geometry
Bunny
Cheongsam
Mask

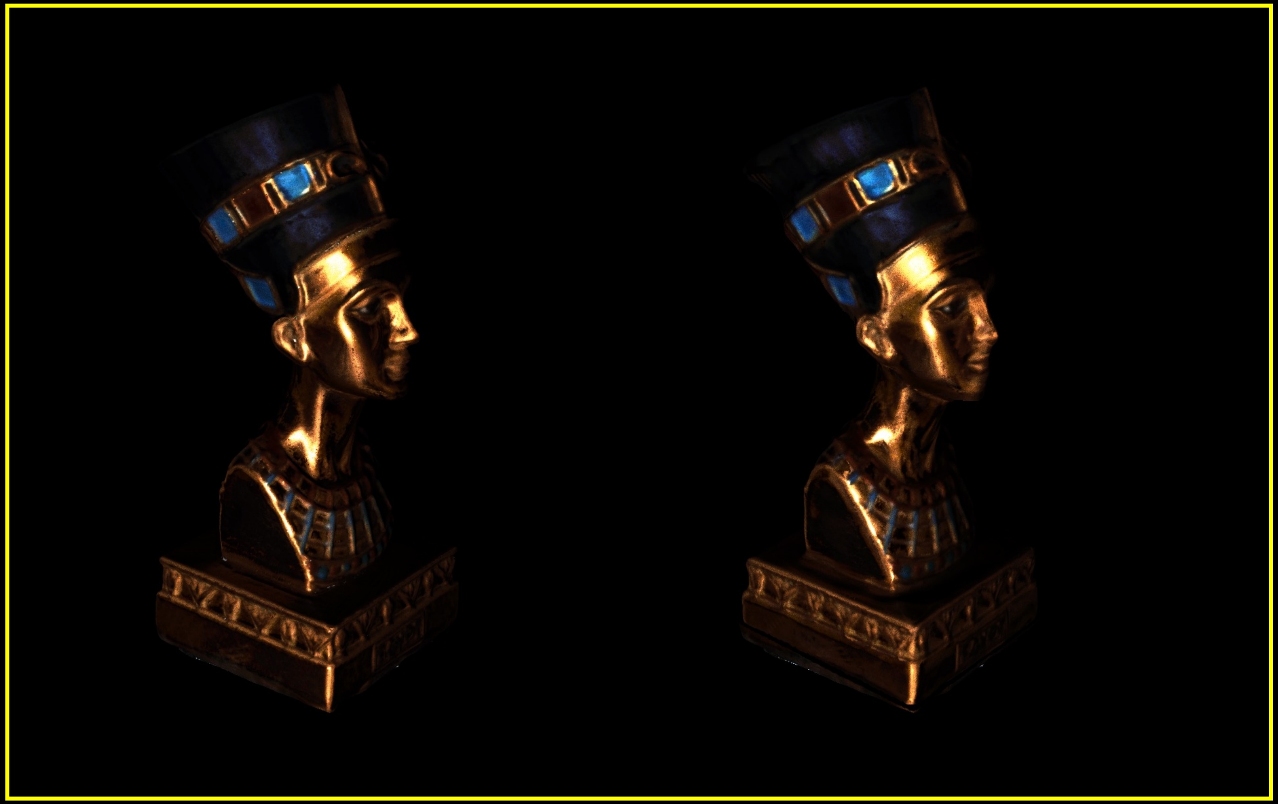
Validation Results

SSIM=0.89 SSIM=0.90 SSIM=0.93 SSIM=0.91 SSIM=0.93 SSIM=0.88 SSIM=0.90

Comparisons

Photo

Ours



Scanned Shape
+
Optimized Pattern

Shape from [Nam et al. 2018]
+
Optimized Pattern

[Nam et al. 2018]

Comparisons

Ours

High-End Lightstage
[Kang et al. 2019]

Evaluations

Repeatability

Scan #1

Scan #2

Scan #1

Scan #2

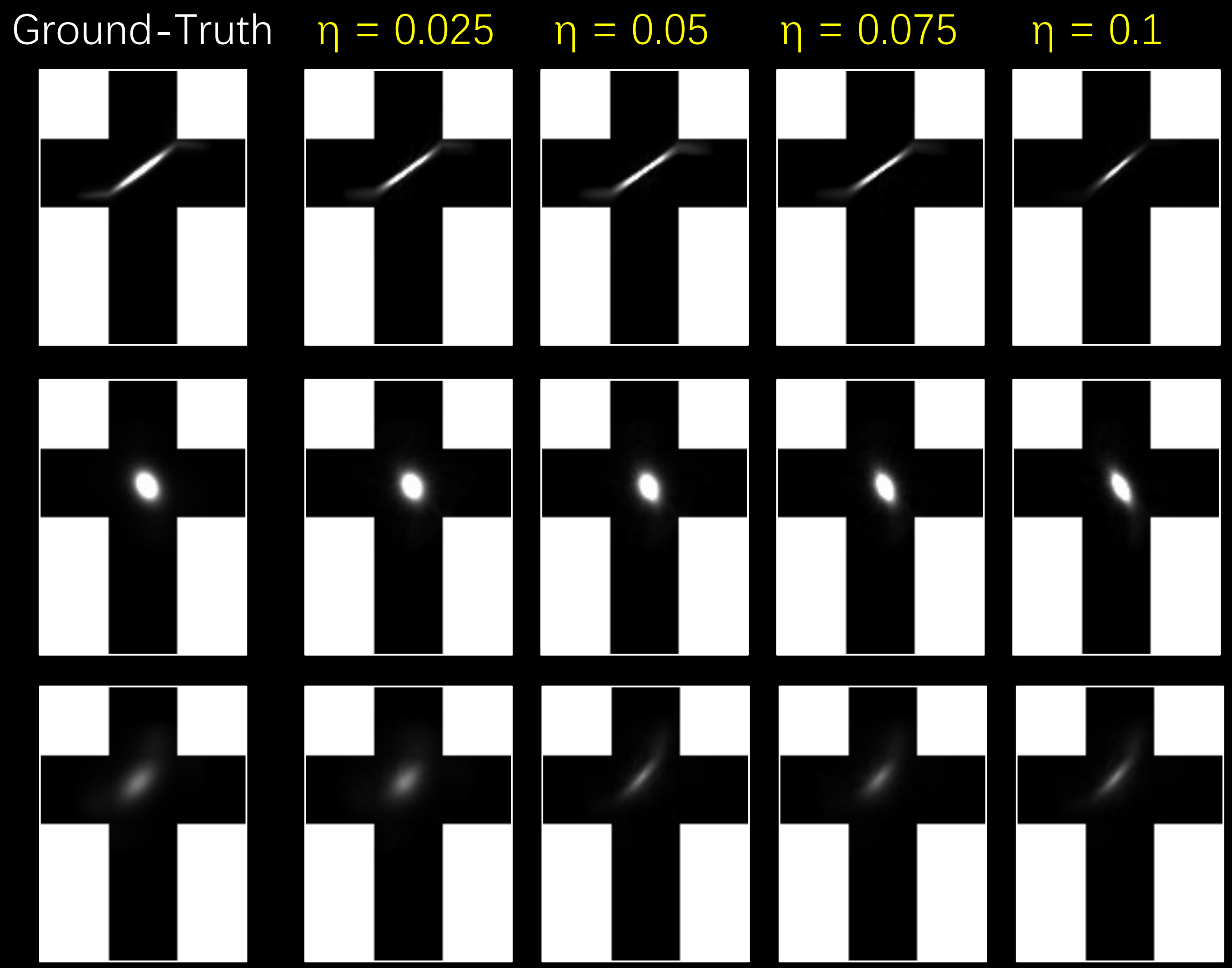
Impact of Geometric Quality

High Quality Mesh from 3D Scanner

Filtered Mesh from 3D Scanner

Mesh from COLMAP

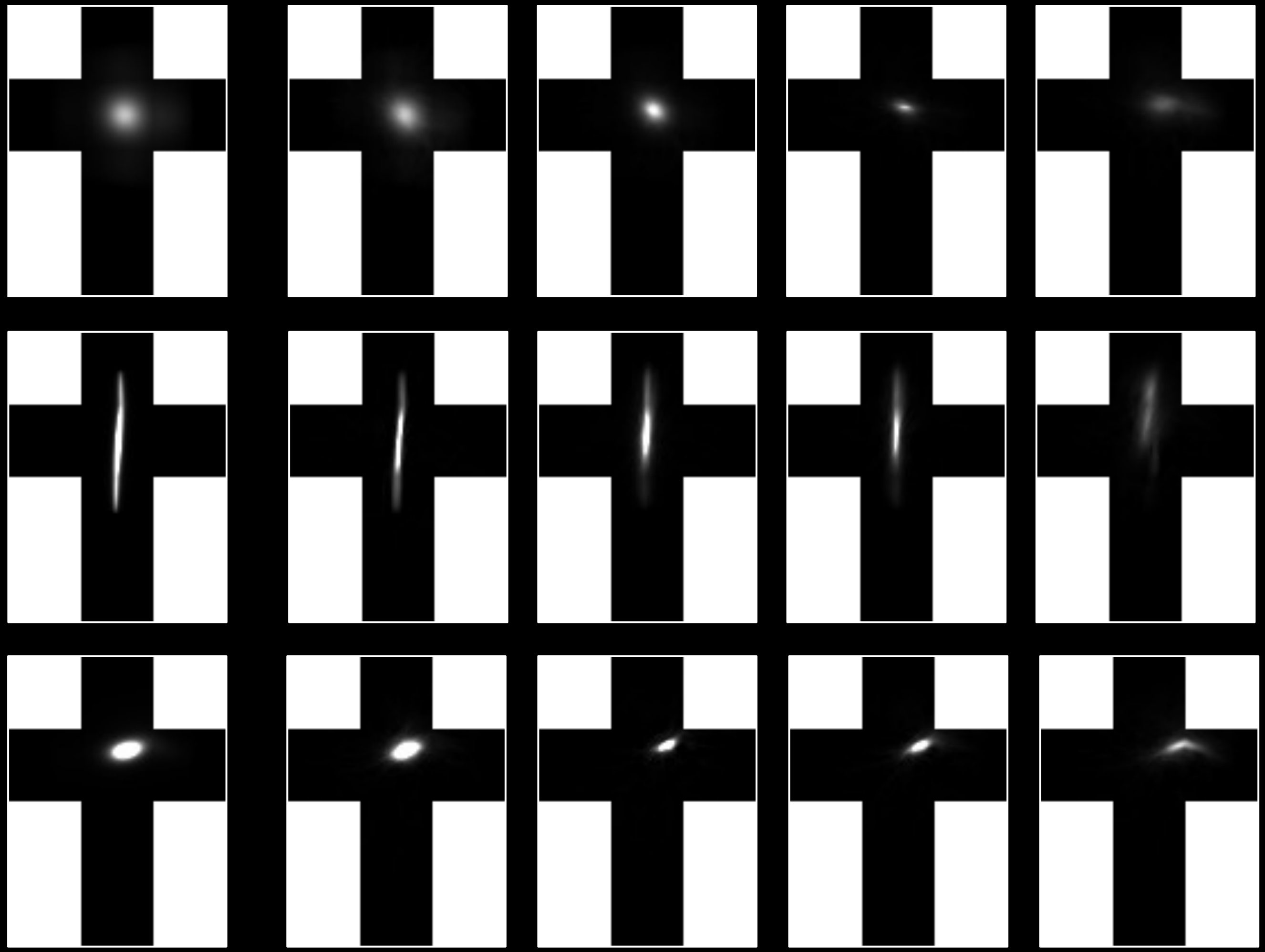
Impact of Camera Pose Error



η: Magnitude of the Camera Pose Error

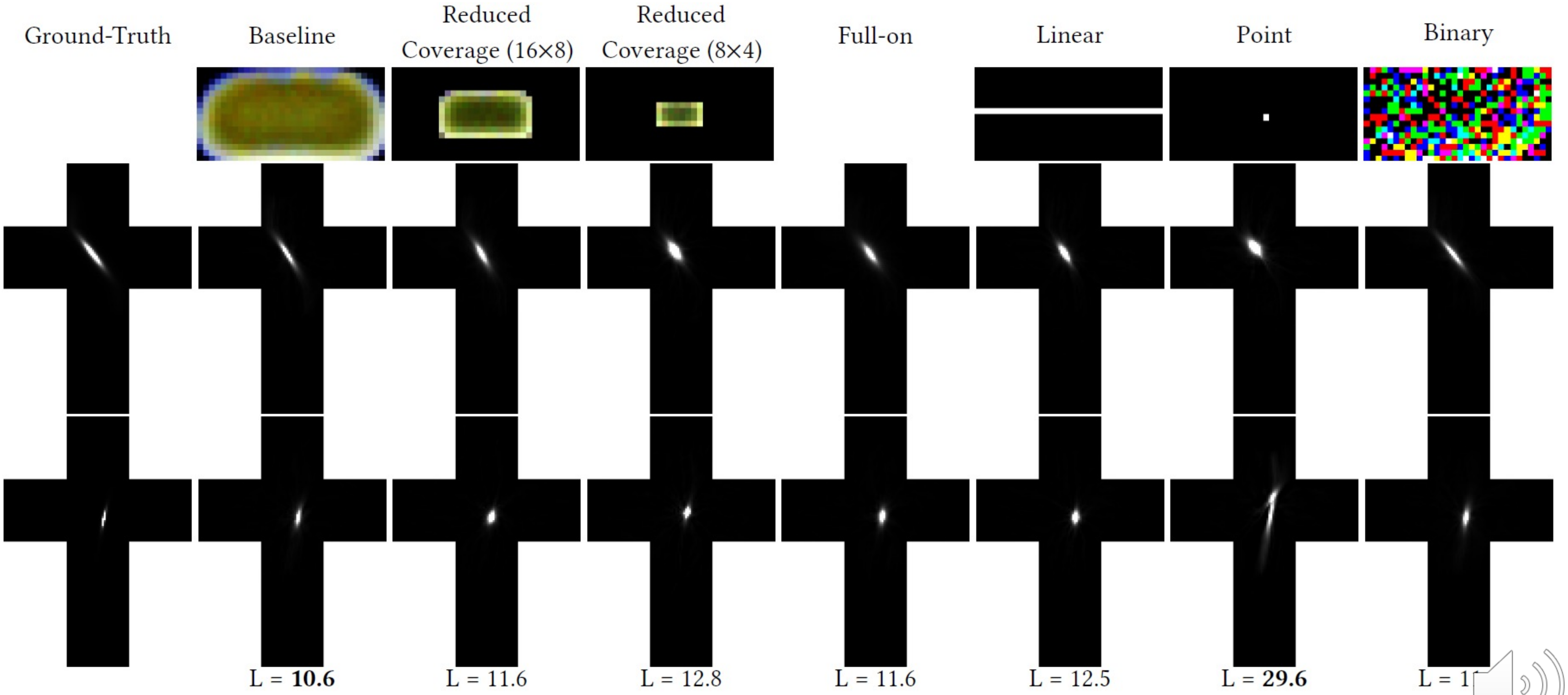
Impact of Specular Highlight Coverage

Ground-Truth ξ = 1.0 ξ = 0.8 ξ = 0.6 ξ = 0.4

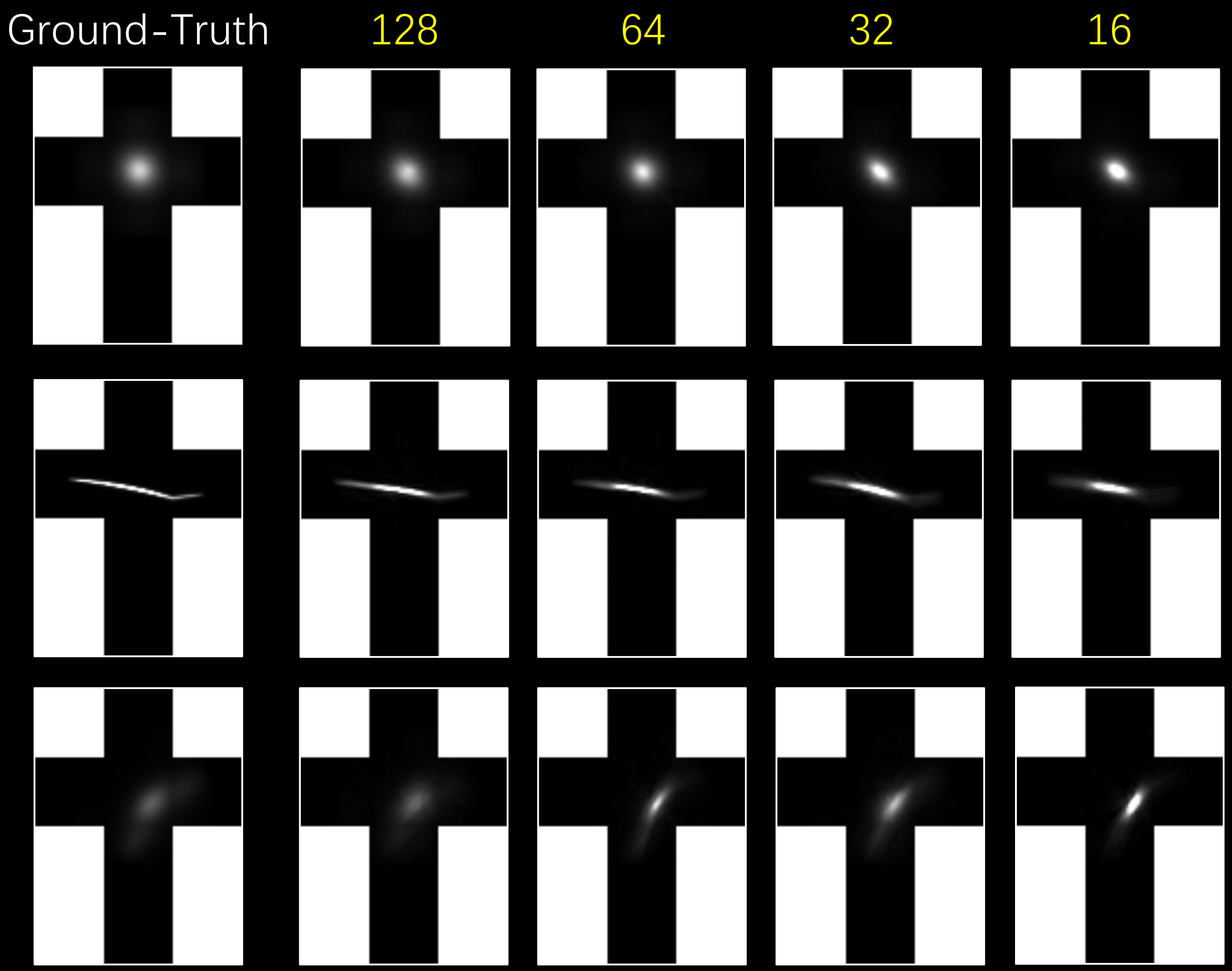


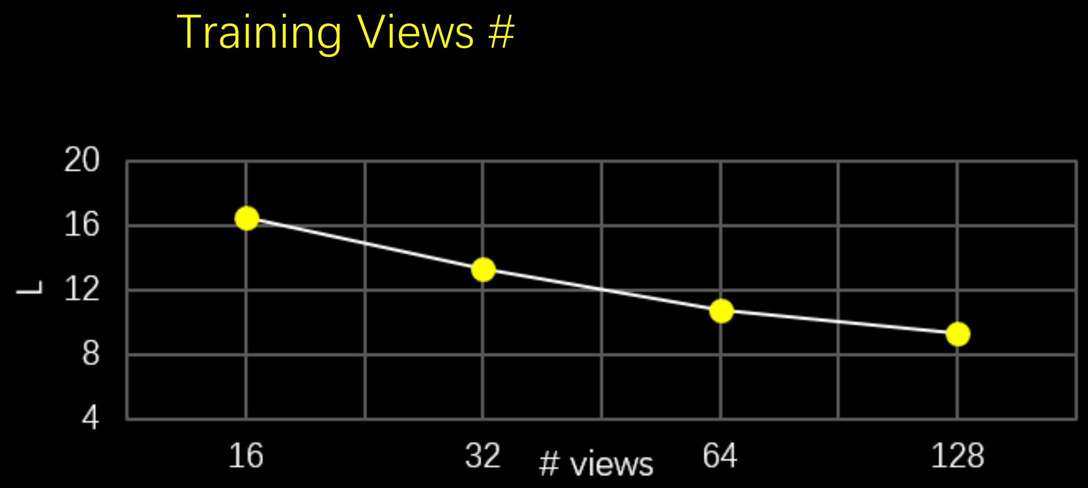
ξ: Reject a View When
Half Vector Satisfies (h · n) > ξ

Impact of Lighting Patterns

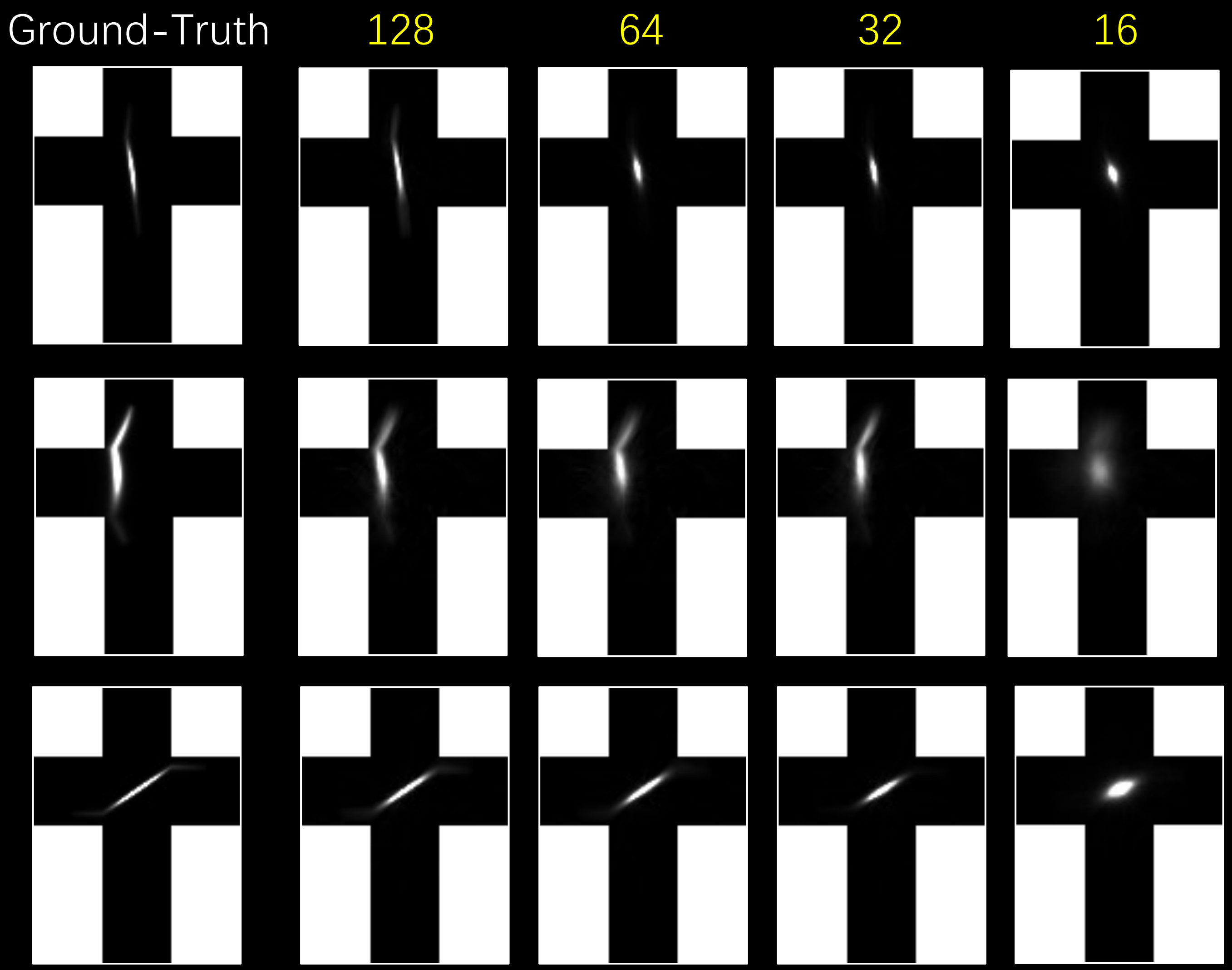


Impact of Training Views





Impact of Test View



Test Views #

Limitations

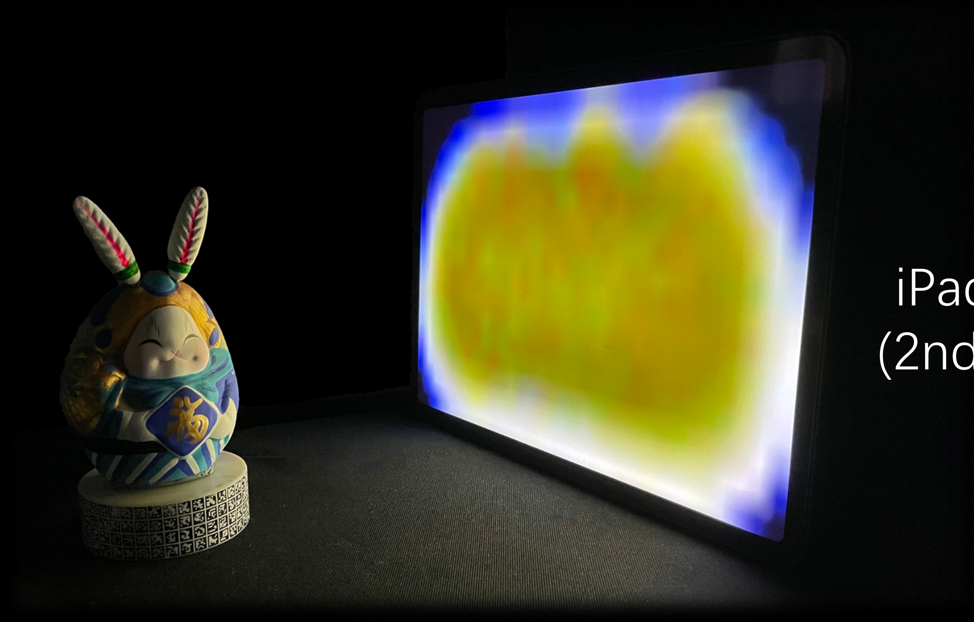
- No Consideration for Global Illumination
- Need a Relatively Precise 3D Shape
- Cannot Recover Appearance Substantially Deviated from Training Samples

Conclusions

- Differentiable Framework for High-quality Scanning of Anisotropic Appearance
  - Neural Trace Photography
- Automatically Learns
  - Lighting Condition
  - Measurements => Reflectance
- Adapts to Various Factors
  - Point/Linear/Area Light
  - Setup's Geometry

Future Work

- Extend to a Similar Device

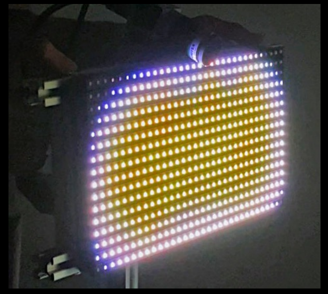


iPad Pro
(2nd Gen)

- Unified Neural Scanner for Shape + Reflectance

Geometry Scanner

+



Appearance Scanner

= ?

Acknowledgements

- Anonymous Reviewers
- Minyi Gu, Yaxin Yu, Zimin Chen, Lijian Ge (Zhejiang University)
- Yang Li (National Museum of China)
- Giljoo Nam, Min H. Kim (KAIST)
- Yue Dong (MSRA)
- Yiruo Zhao

- National Key Research & Development Program of China (2018YFB1004300)
- National Science Foundation of China (61772457, 62022072 & 61890954)

SIGGRAPH 2021
VIRTUAL 9-13 AUGUST

Thank you/謝謝