Learning Efficient Illumination Multiplexing for Joint Capture of Reflectance and Shape

Kaizhang Kang, Cihui Xie, Cheng'an He, Mingqi Yi,
Minyi Gu, Zimin Chen, Kun Zhou and Hongzhi Wu

State Key Lab of CAD&CG, Zhejiang University / Yale University /
ZJU-FaceUnity Joint Lab of Intelligent Graphics

SIGGRAPH ASIA 2019 BRISBANE

Introduction

- Realistic Digital Models are Important

Culture Heritage

e-Commerce

Visual Effects

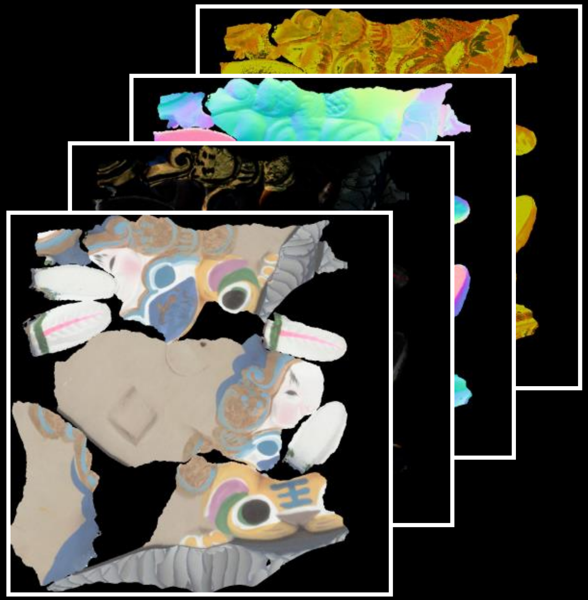
https://www.coinsshopy.com/product/classic-brand-shoes-women-casual-pointed-toe-black-oxford-shoes-for-women-flats-comfortable-slip-on-women-shoes/

Introduction

- Realistic Digital Models are Important

3D Mesh

+



6D SVBRDF
(Location, Lighting & View Directions)

=

Digital Model

Introduction

- Realistic Digital Models are Important
- Acquisition of Physical Objects is Crucial in Graphics/Vision

Introduction

- Realistic Digital Models are Important
- Acquisition of Physical Objects is Crucial in Graphics/Vision
- Efficient, Joint Capture of Reflectance & Shape is Challenging
    1. High-dimensional Unknowns
    2. Reflectance & Shape Tightly Coupled in Measurements
    3. Limited Number of Samples in Practice

Introduction

- Realistic Digital Models are Important
- Acquisition of Physical Objects is Crucial in Graphics/Vision
- Efficient Acquisition of Reflectance & Shape is Challenging
- Our Goal
  - Optimize Physical Acquisition Efficiency in Joint Capture of Reflectance & Shape

Our Framework

- Map Physical Acquisition & Computational Reconstruction to a Deep Neural Network
  - Automatic Optimization of Illumination w.r.t Joint Acquisition Efficiency
  - Breaks Mutual Dependency between Reflectance & Shape

Our Framework

- Map Physical Acquisition & Computational Reconstruction to a Deep Neural Network
- Carefully Designed Network Architecture
  - Shares Information between Reflectance & Shape Estimation
  - Combines Domain-Specific Knowledge with Deep Learning

Our Framework

- Map Physical Acquisition & Computational Reconstruction to a Deep Neural Network
- Carefully Designed Network Architecture
- Flexible / Adaptable
  - Setup’s Geometry
  - Properties of Appearance

Related Work

- Geometry Reconstruction with a Diffuse Assumption
  - Structured Lighting [Scharstein and Szeliski 2003] / Structure-from-Motion [Schonberger et al. 2016]
    - Diffuse-dominant Reflectance Assumption
  - Photometric Stereo [Woodham 1980]
    - Latest Work Limited to Isotropic Reflectance [Ikehata 2018]

Related Work

- Geometry Reconstruction with a Diffuse Assumption
- Reflectance Capture on a Known/Pre-acquired Shape
  - Direct Sampling [Dana et al. 1999; Lawrence et al. 2006]
  - Reflectance Prior [Dong et al. 2010; Aittala et al. 2015; Wu et al. 2016]
  - Illumination Multiplexing [Gardner et al. 2003; Ghosh et al. 2009; Aittala et al. 2013; Kang et al. 2018]

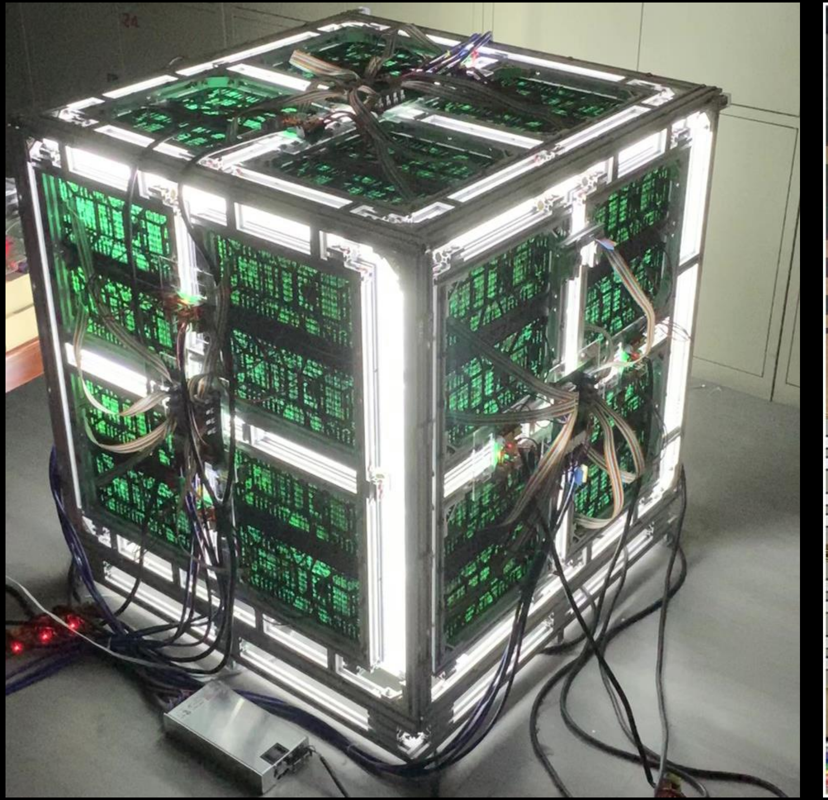
Related Work

- Geometry Reconstruction with a Diffuse Assumption
- Reflectance Capture on a Known/Pre-acquired Shape
- Joint Acquisition of Reflectance & Shape
    - Reflectance Prior [Holroyd et al. 2010; Zhou et al. 2013; Nam et al. 2018]
    - Illumination Prior [Tunwattanapong et al. 2013; Xia et al. 2016]
    - Alternating Optimization [Nam et al. 2018]
    - Physical Efficiency not Optimized

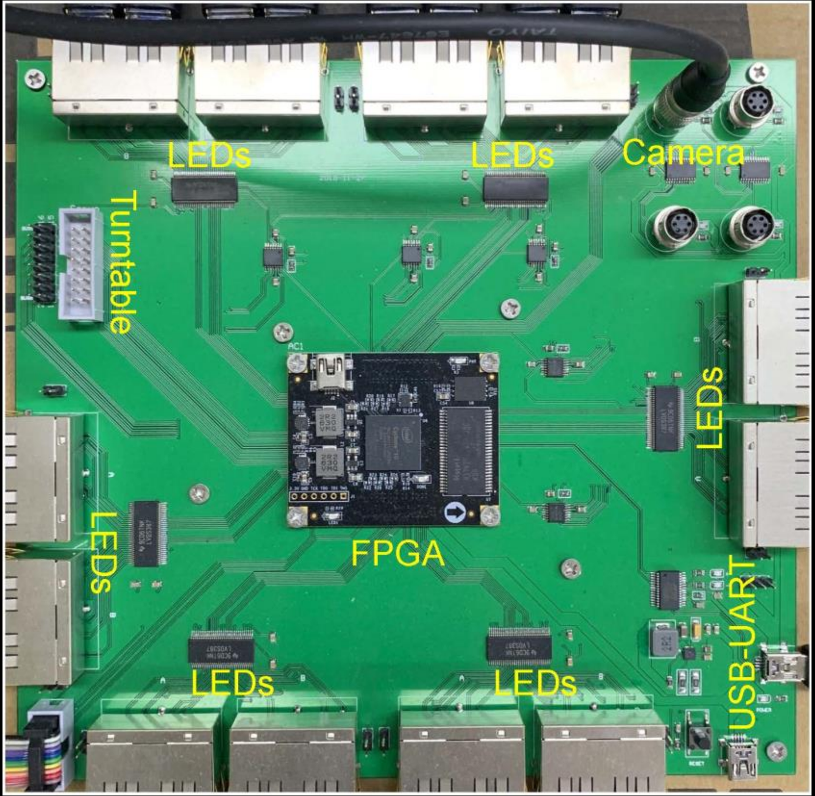
Related Work

- Geometry Reconstruction with a Diffuse Assumption
- Reflectance Capture on a Known/Pre-acquired Shape
- Joint Acquisition of Reflectance & Shape
- Deep-Learning-Assisted Modeling
  - Reflectance Modeling [Li et al. 2017; Deschaintre et al. 2018]
  - Shape Modeling [Kendall et al. 2017; Yao et al. 2018]
  - Joint Modeling [Li et al. 2018]
  - Focus on Shape / Reflectance Reconstruction from Highly Sparse Input
  - Physical Acquisition Process Not Optimized

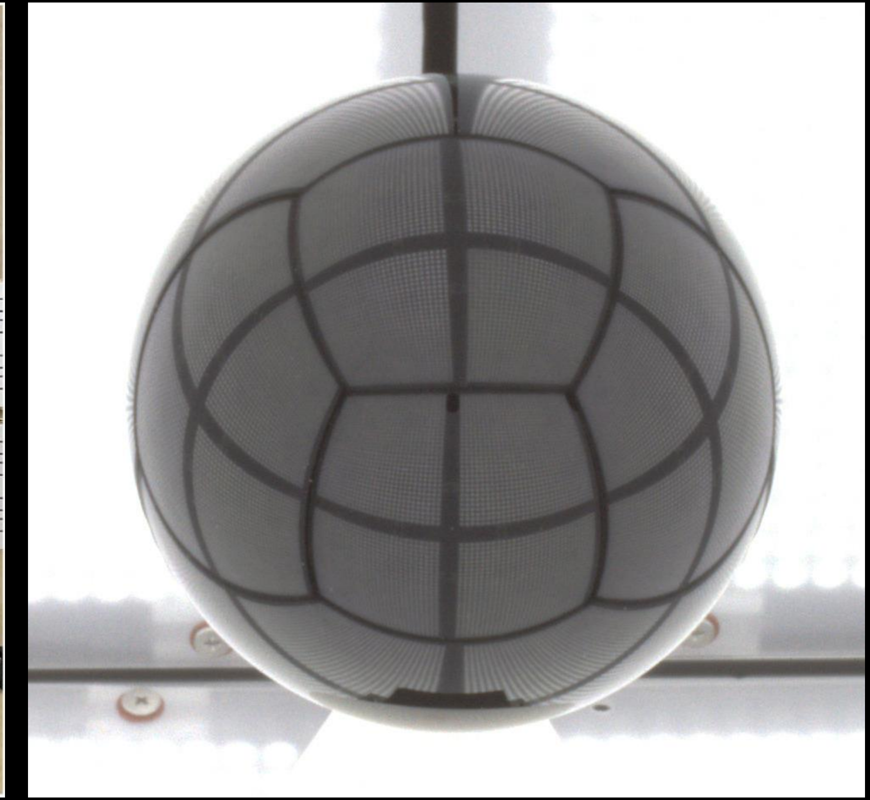
Hardware Prototype



Exterior

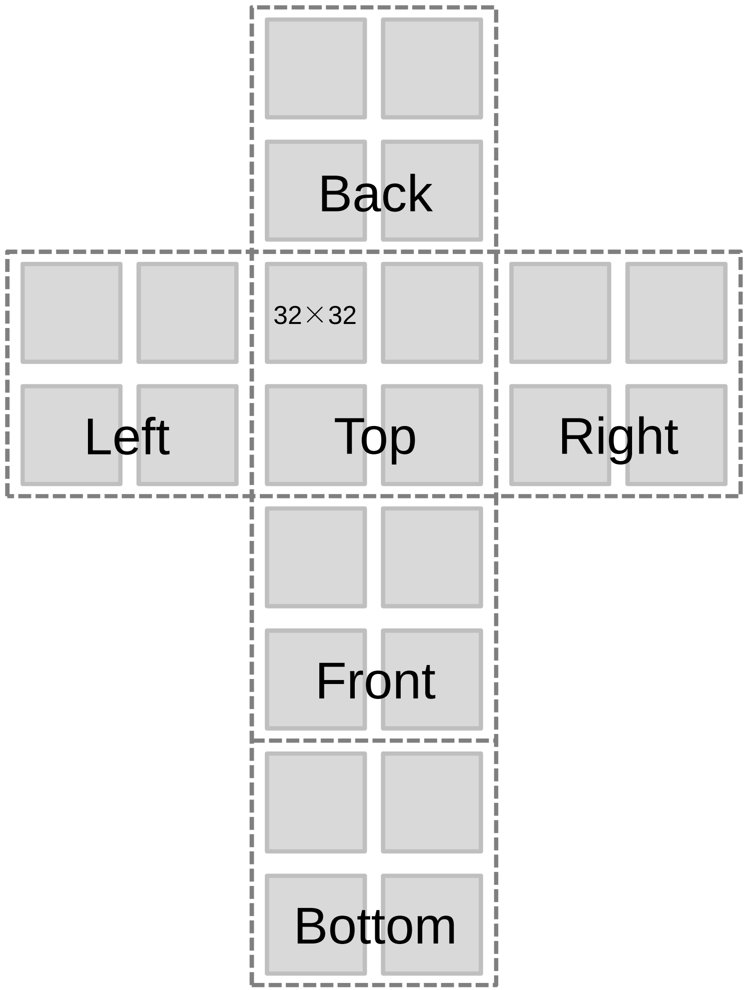


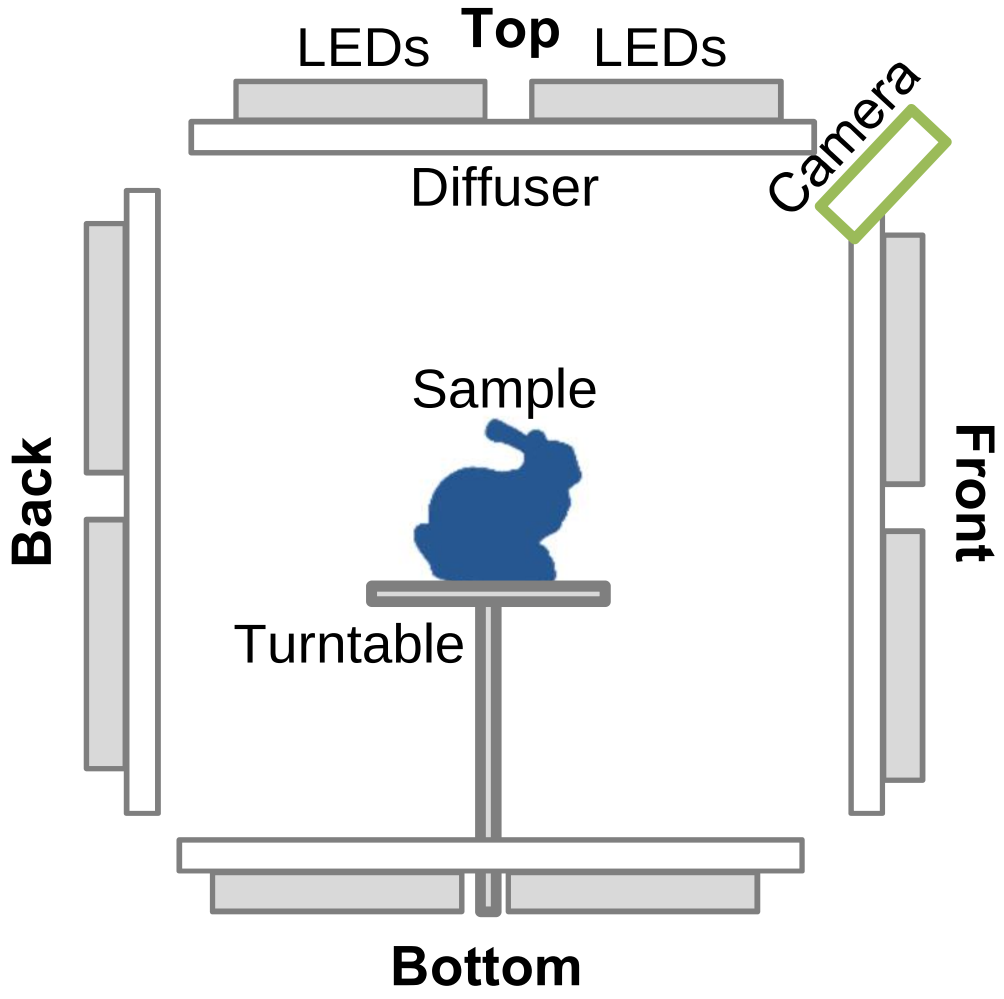
Main Circuit Board



Lit Sphere

Hardware Prototype



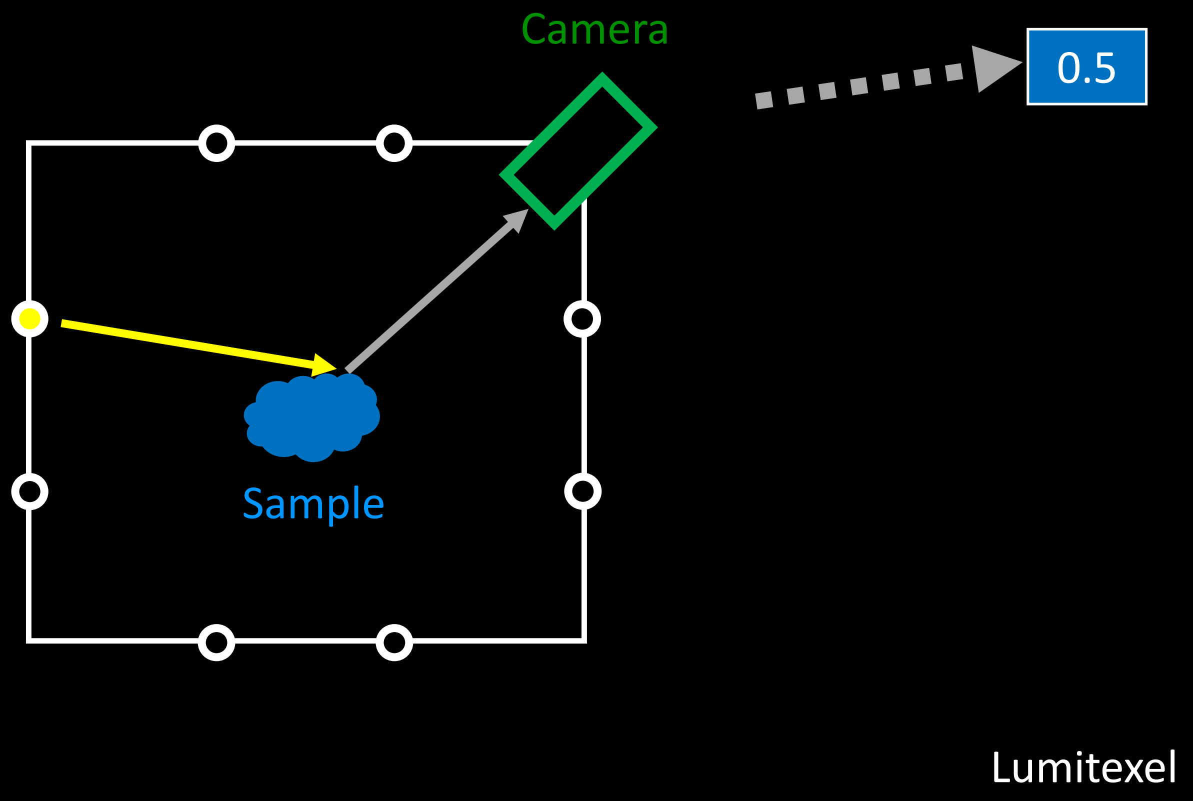


Hardware Prototype

- 80cm x 80cm x 80cm
- Single Camera
- 24,576 LEDs
- 20,000+ FPS for Binary Lighting Patterns

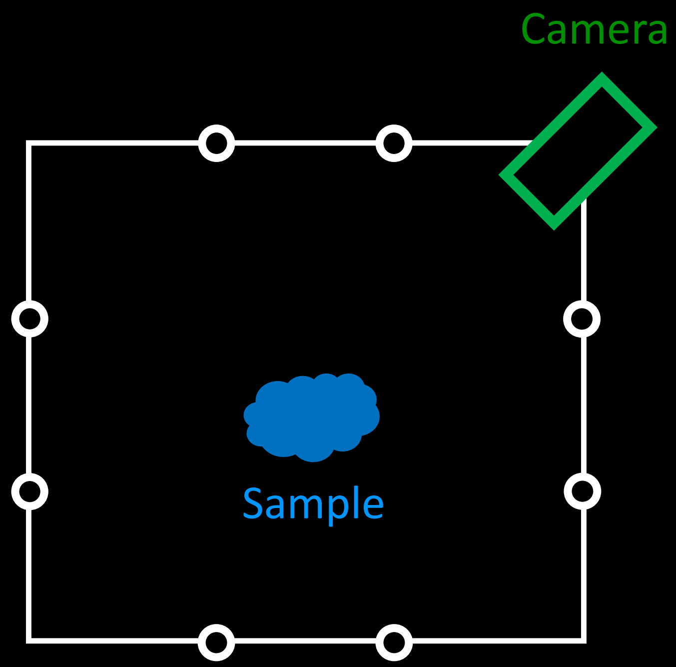
- High-precision Control / Synchronization via FPGA

Lumitexel

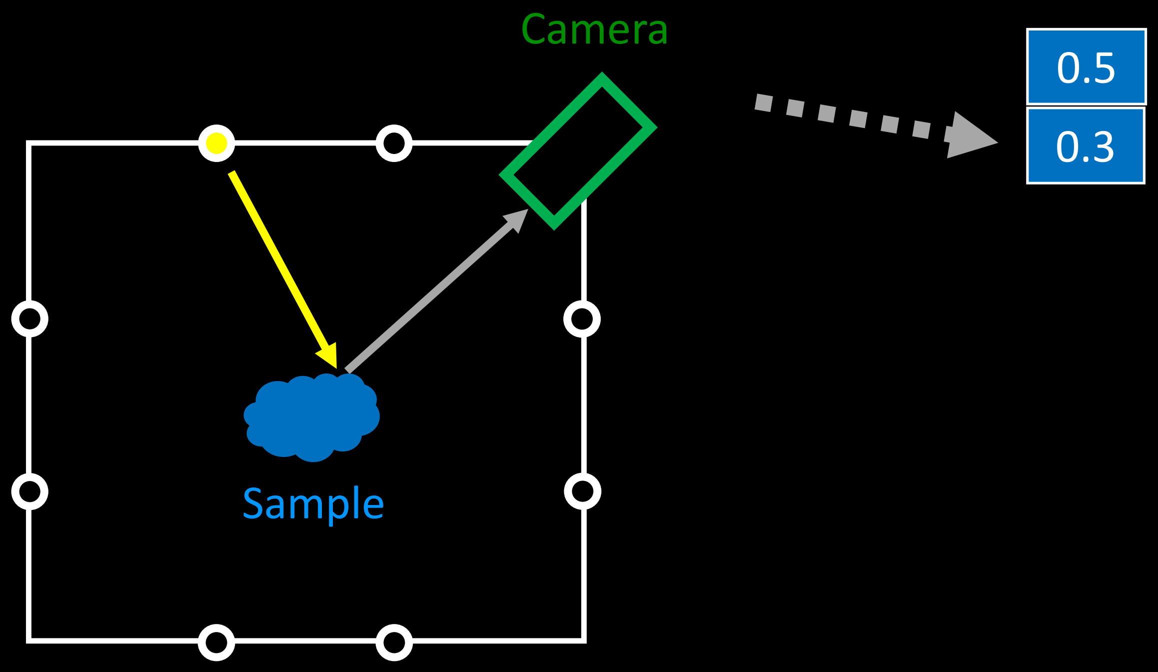


Lumitexel

Lumitexel

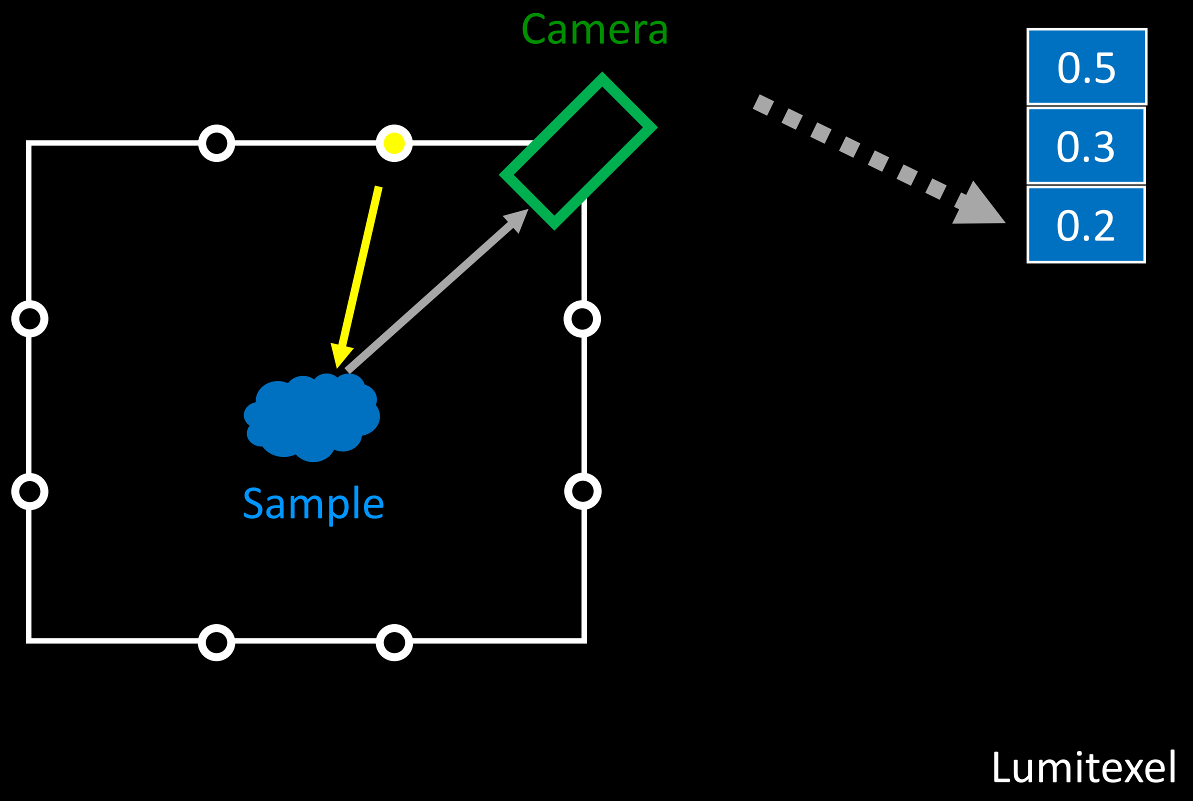


Lumitexel

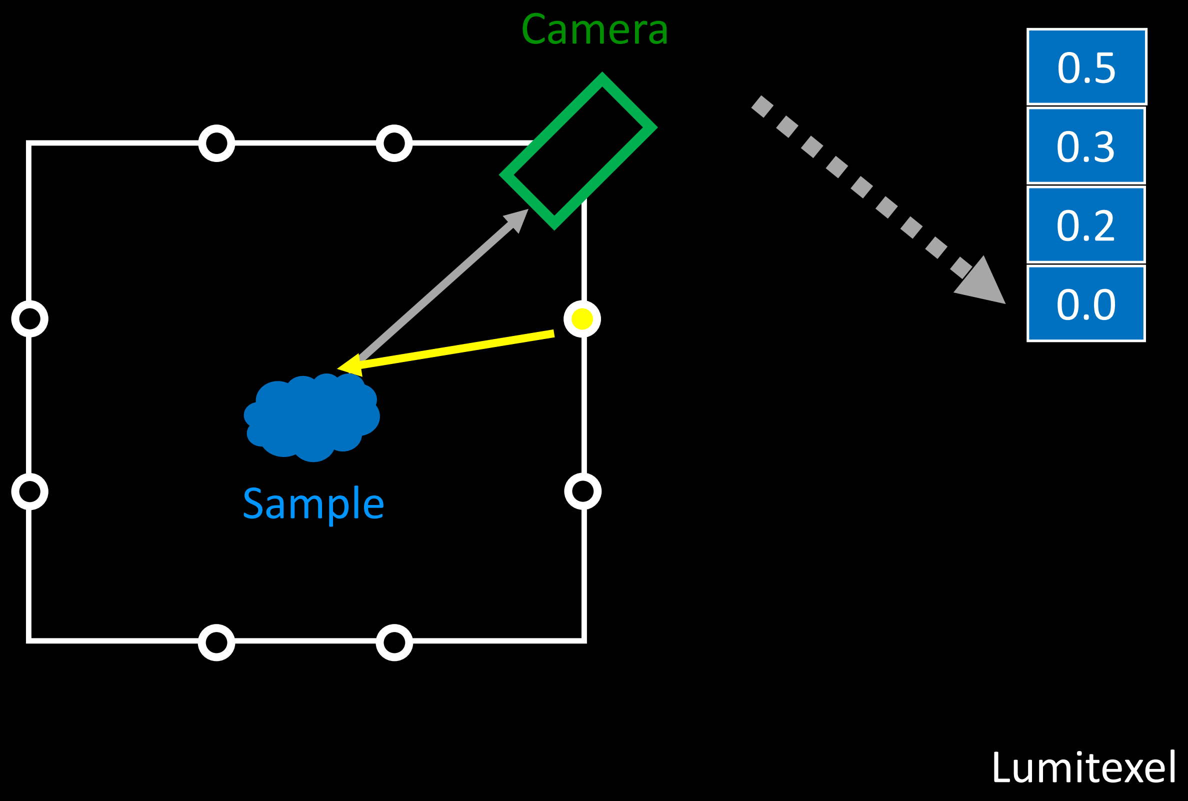


Lumitexel

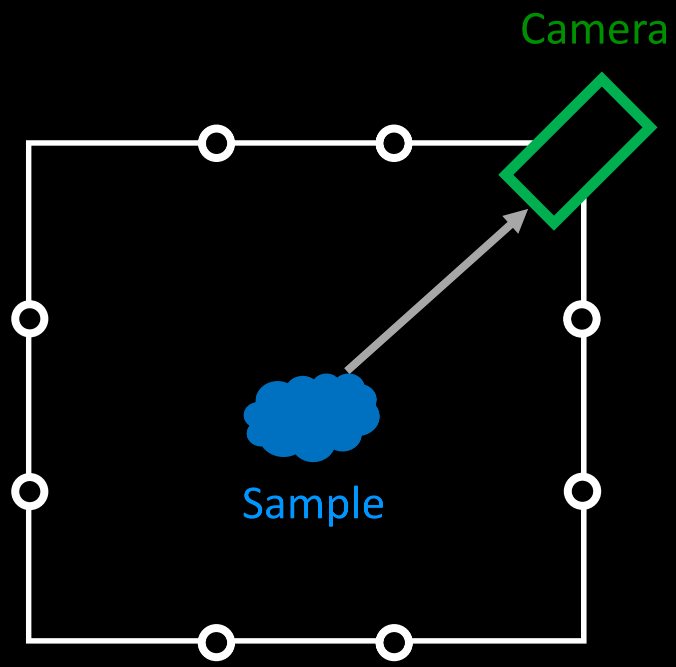
Lumitexel



Lumitexel



Lumitexel



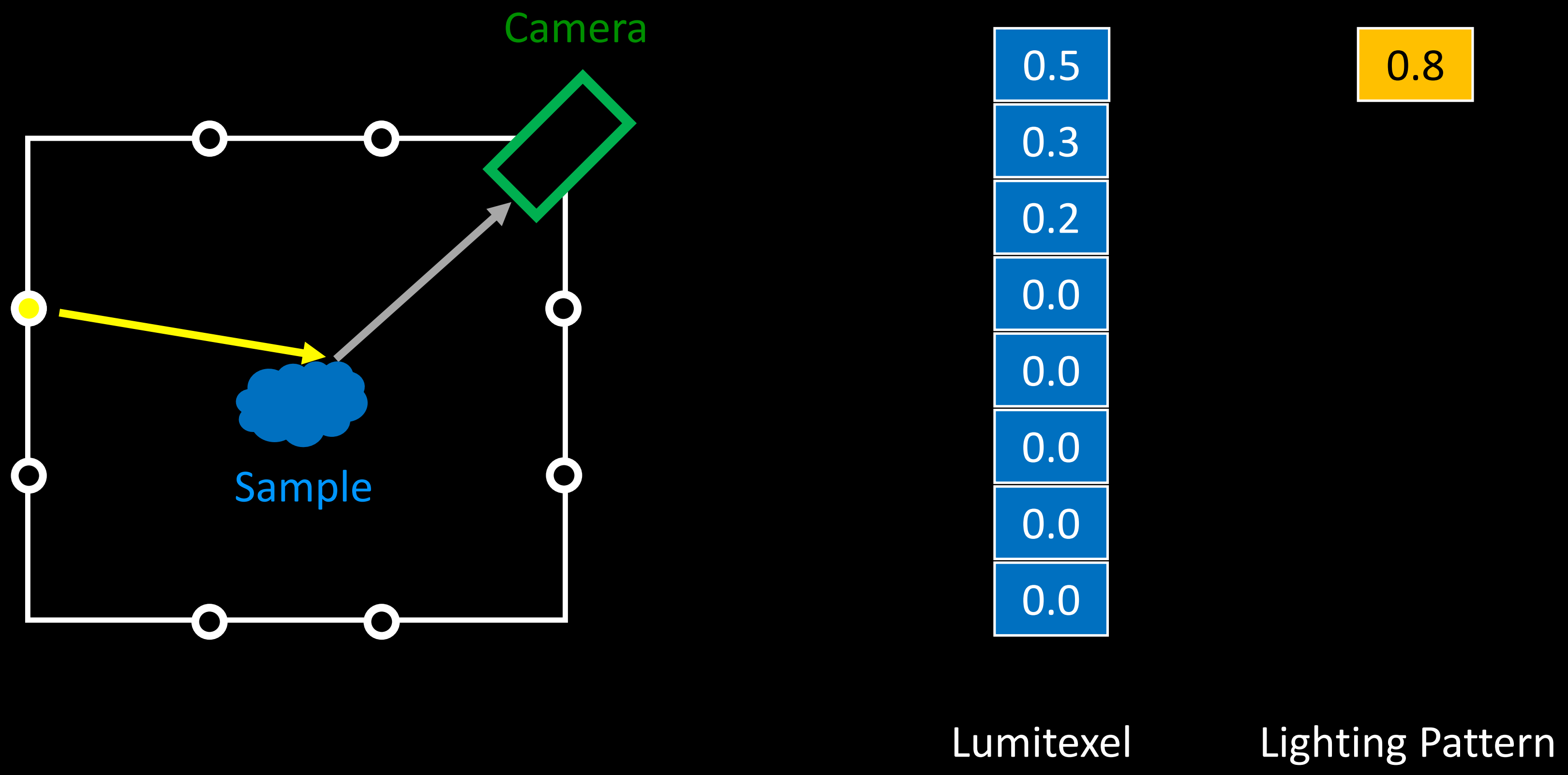
0.5
0.3
0.2
0.0
0.0
0.0
0.0
0.1

Lumitexel

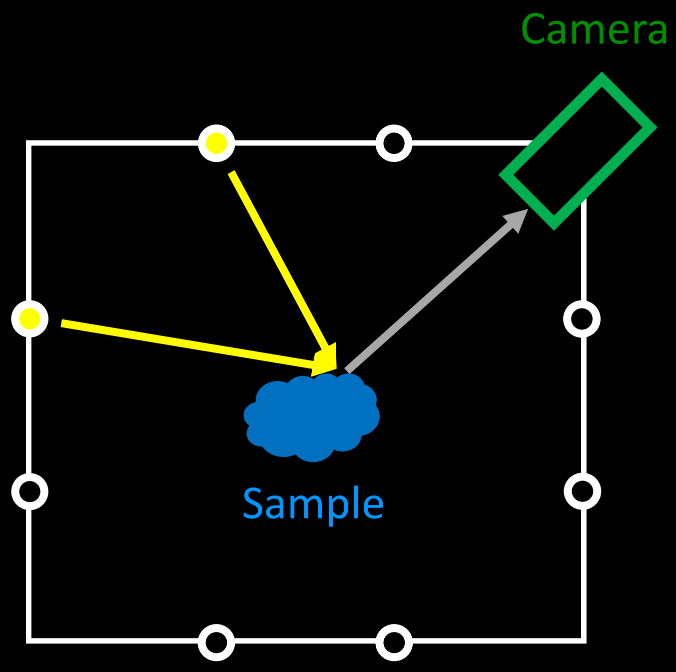
+ Most Informative

- Expensive to Capture

Illumination Multiplexing



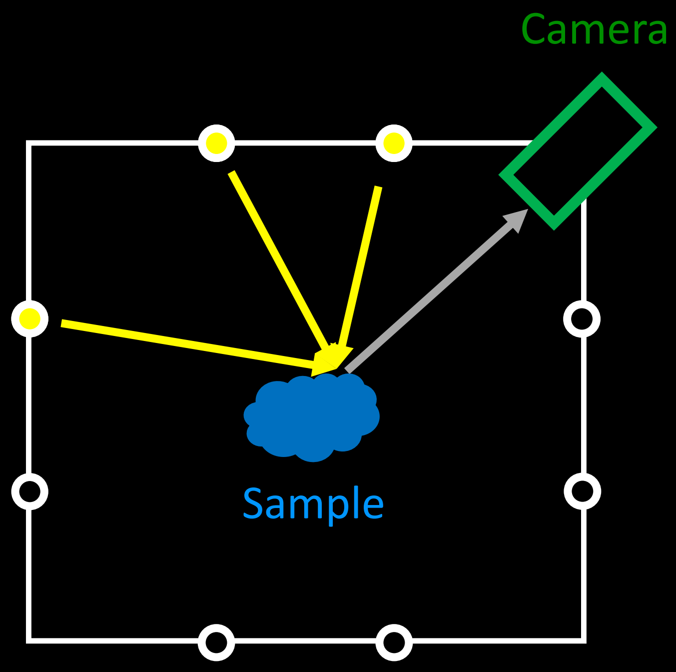
Illumination Multiplexing



Lumitexel
0.5
0.3
0.2
0.0
0.0
0.0
0.0
0.0

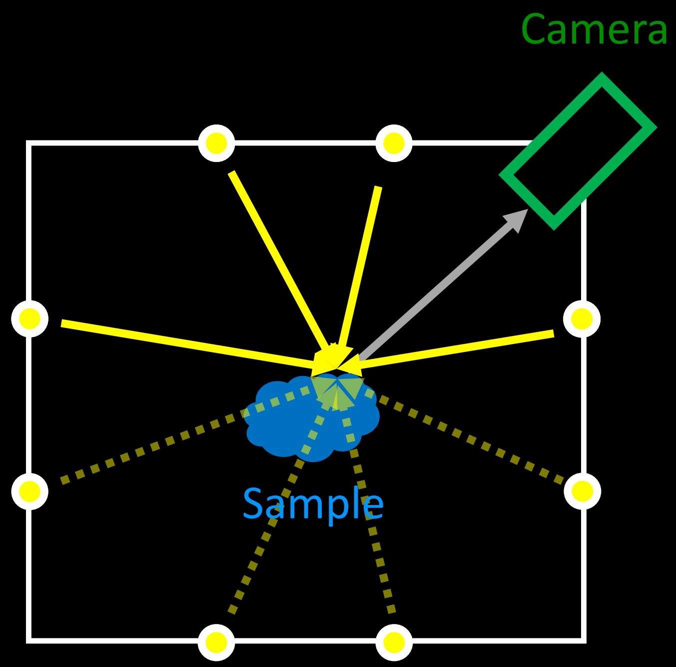
Lighting Pattern
0.8
1.0

Illumination Multiplexing



Lumitexel	Lighting Pattern
0.5	0.8
0.3	1.0
0.2	0.5
0.0	
0.0	
0.0	
0.0	
0.0	

Illumination Multiplexing

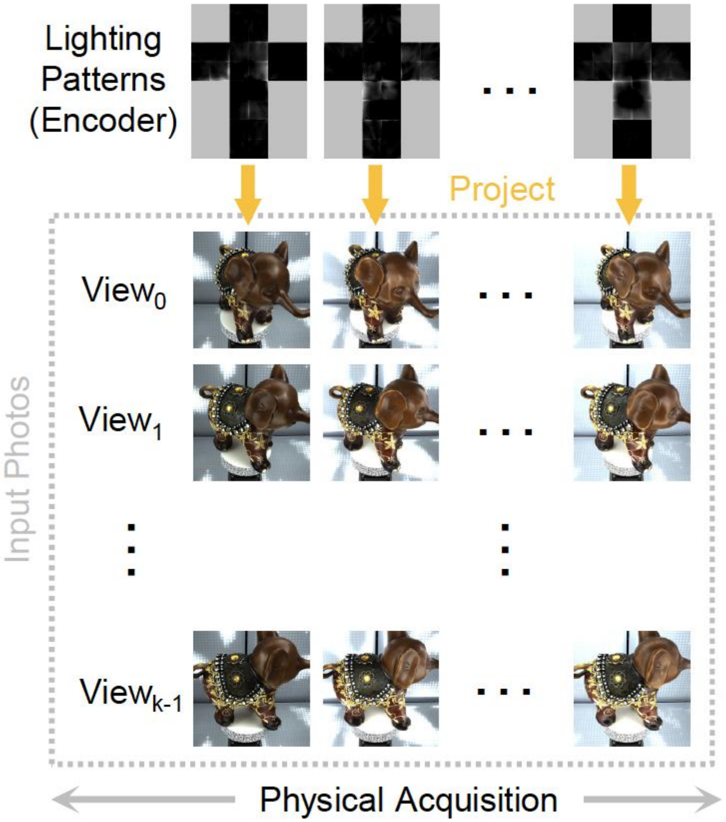


Lumitexel		Lighting Pattern		Measured Radiance
0.5		0.8		
0.3		1.0		
0.2		0.5		
0.0		0.0		
0.0	*	0.4	=	0.8
0.0		1.0		
0.0		0.3		
0.0		0.9		

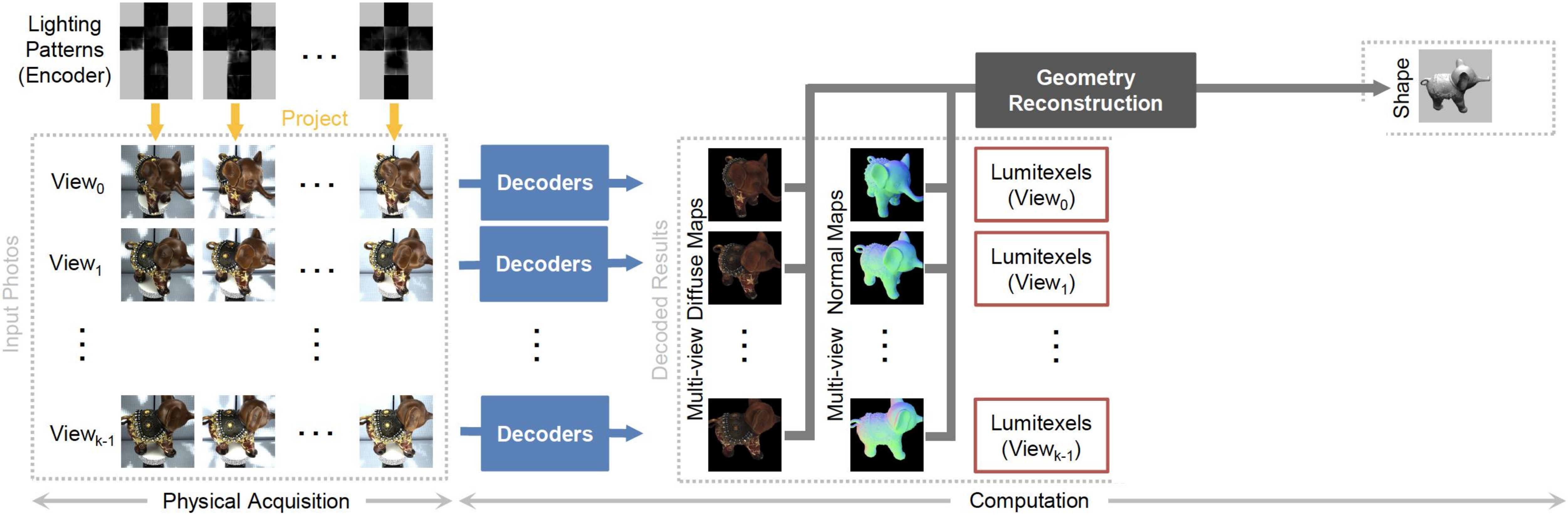
Illumination Multiplexing

- What are Optimal Lighting Patterns for Efficient, Joint Capture of Reflectance & Shape?
- How to Reconstruct Reflectance & Shape from Measurements under Such Patterns?

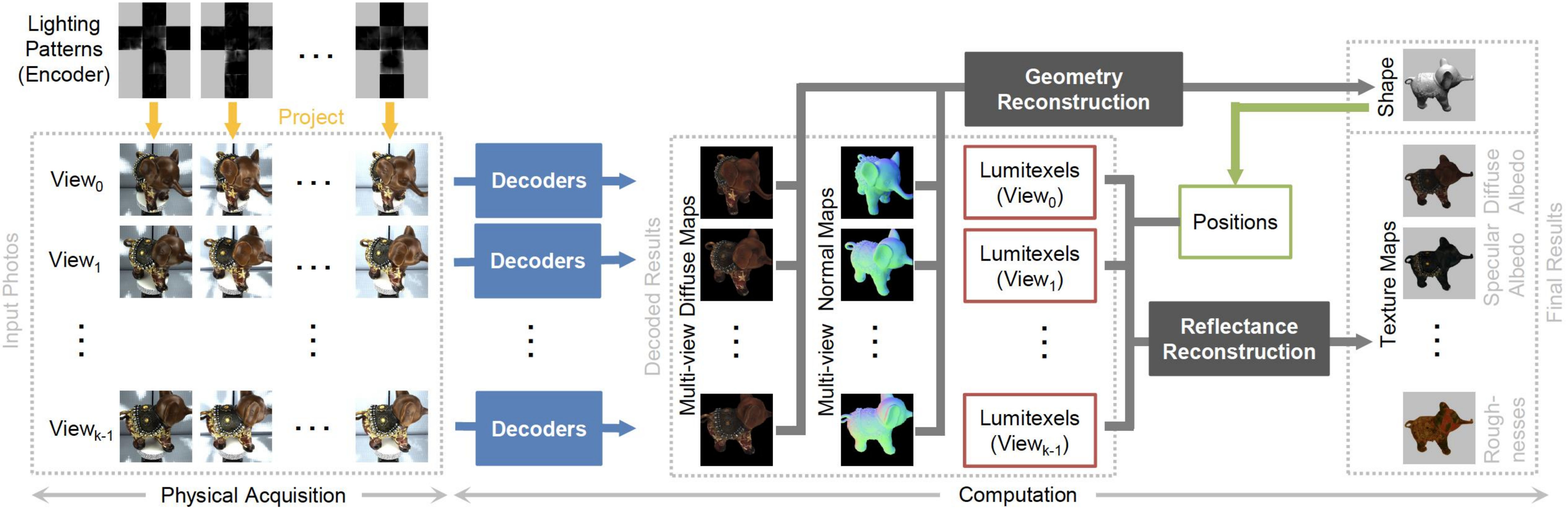
Our Pipeline



Our Pipeline

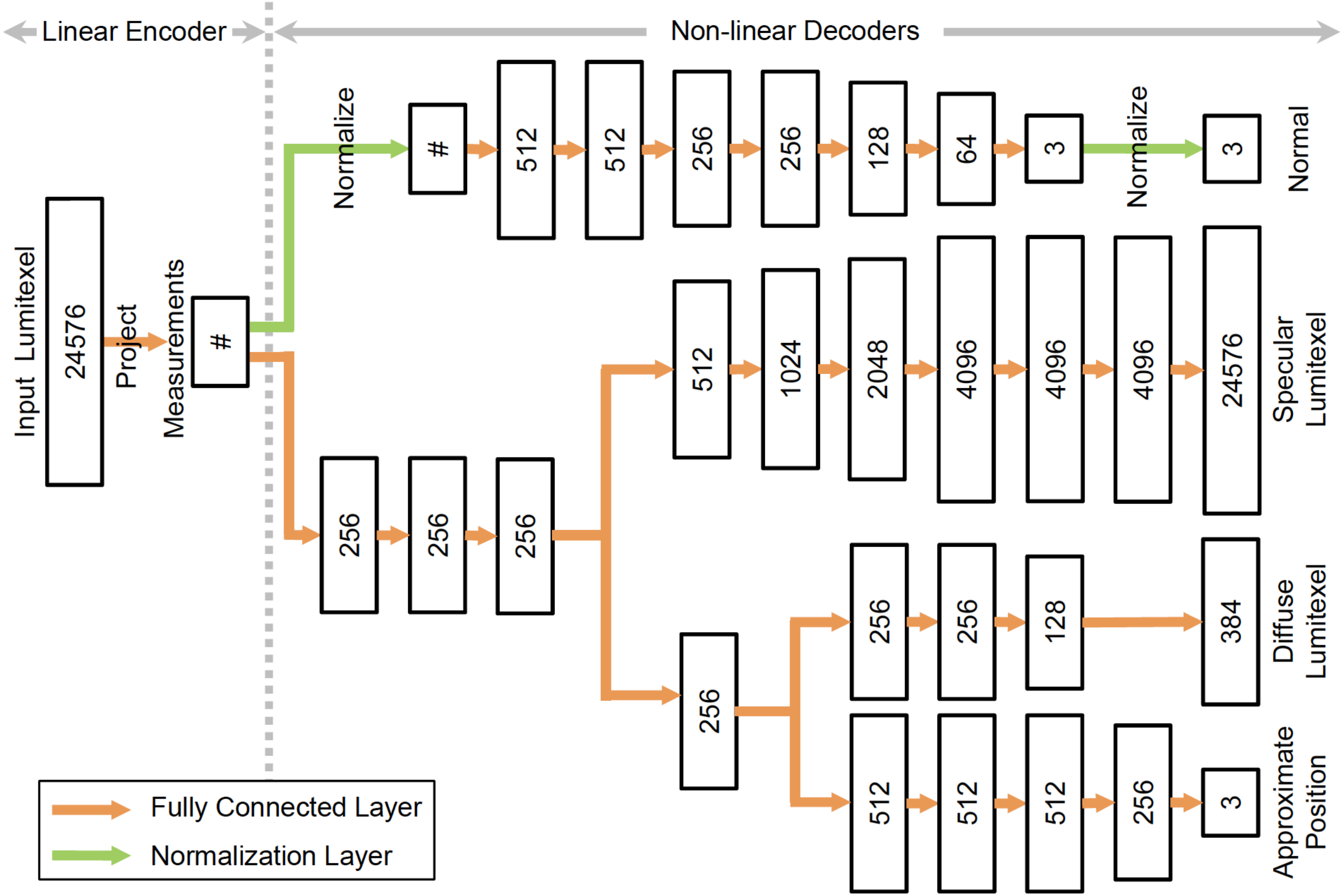


Our Pipeline



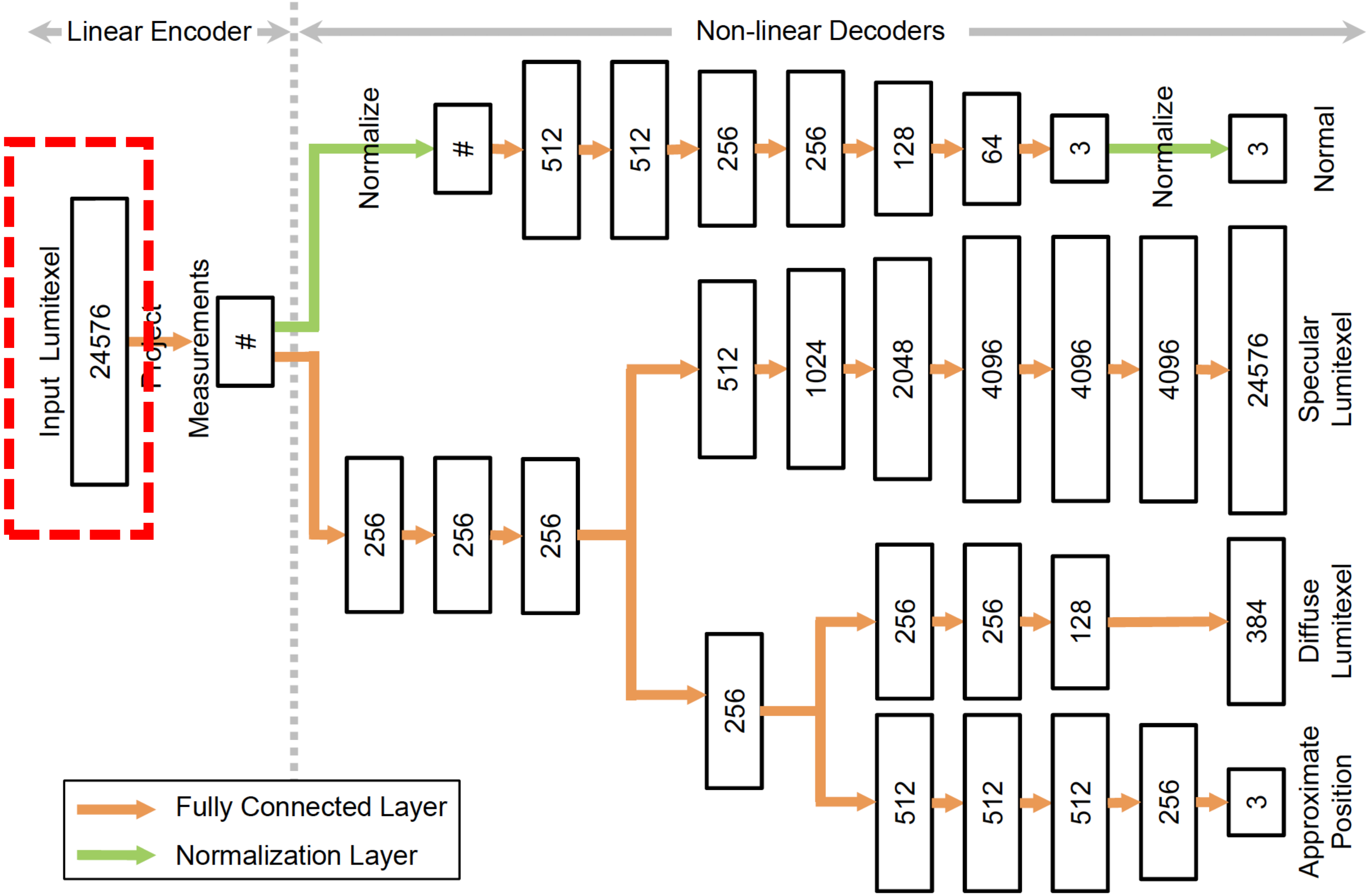
Our Network

- 1 Encoder
  - Physical Capture
- 4 Decoders
  - Computational Reconstruction
- Asymmetric
- Mixed-Domain
- Per-Pixel



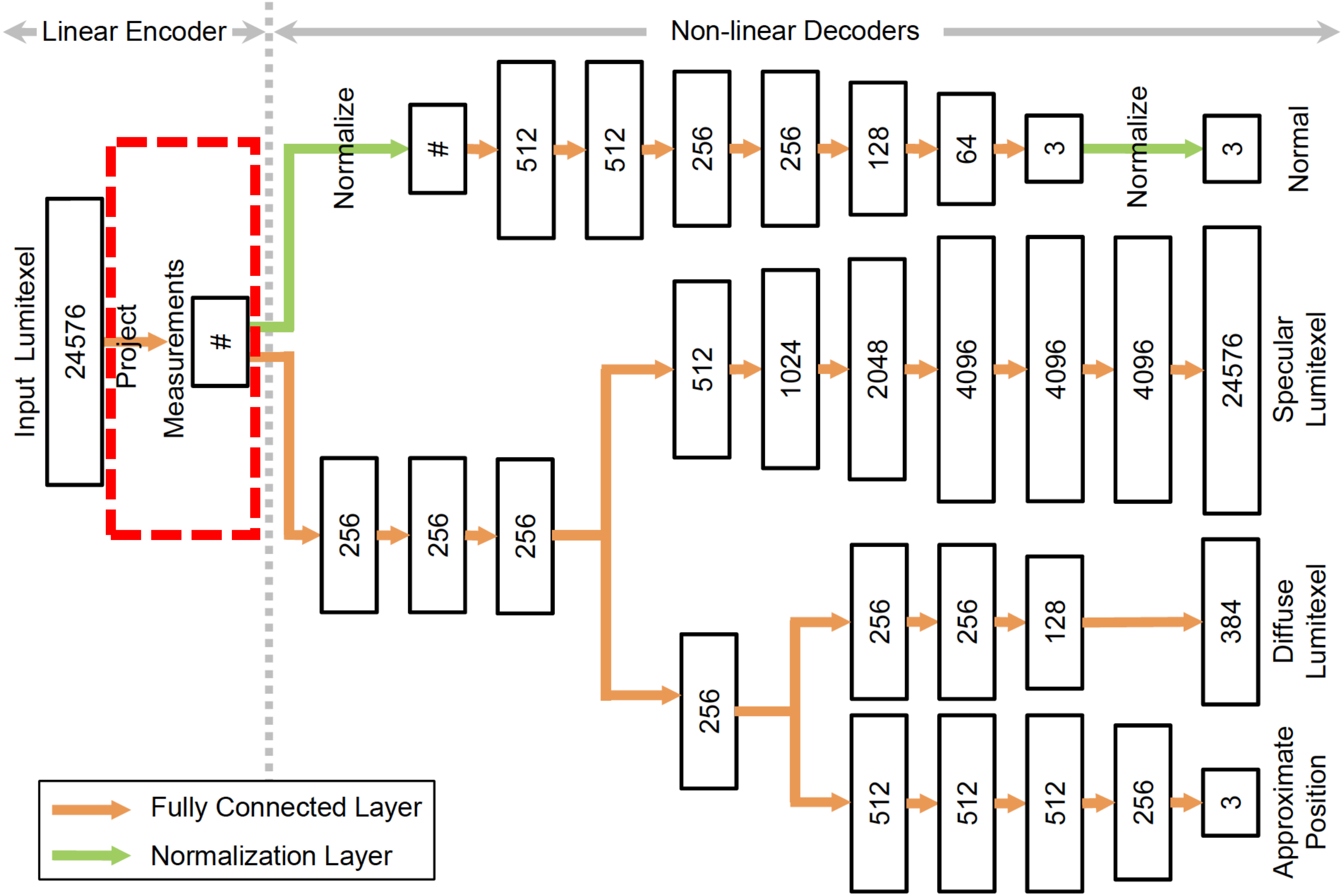
Our Network

- 1 Encoder
  - Physical Capture
- 4 Decoders
  - Computational Reconstruction
- Asymmetric
- Mixed-Domain
- Per-Pixel



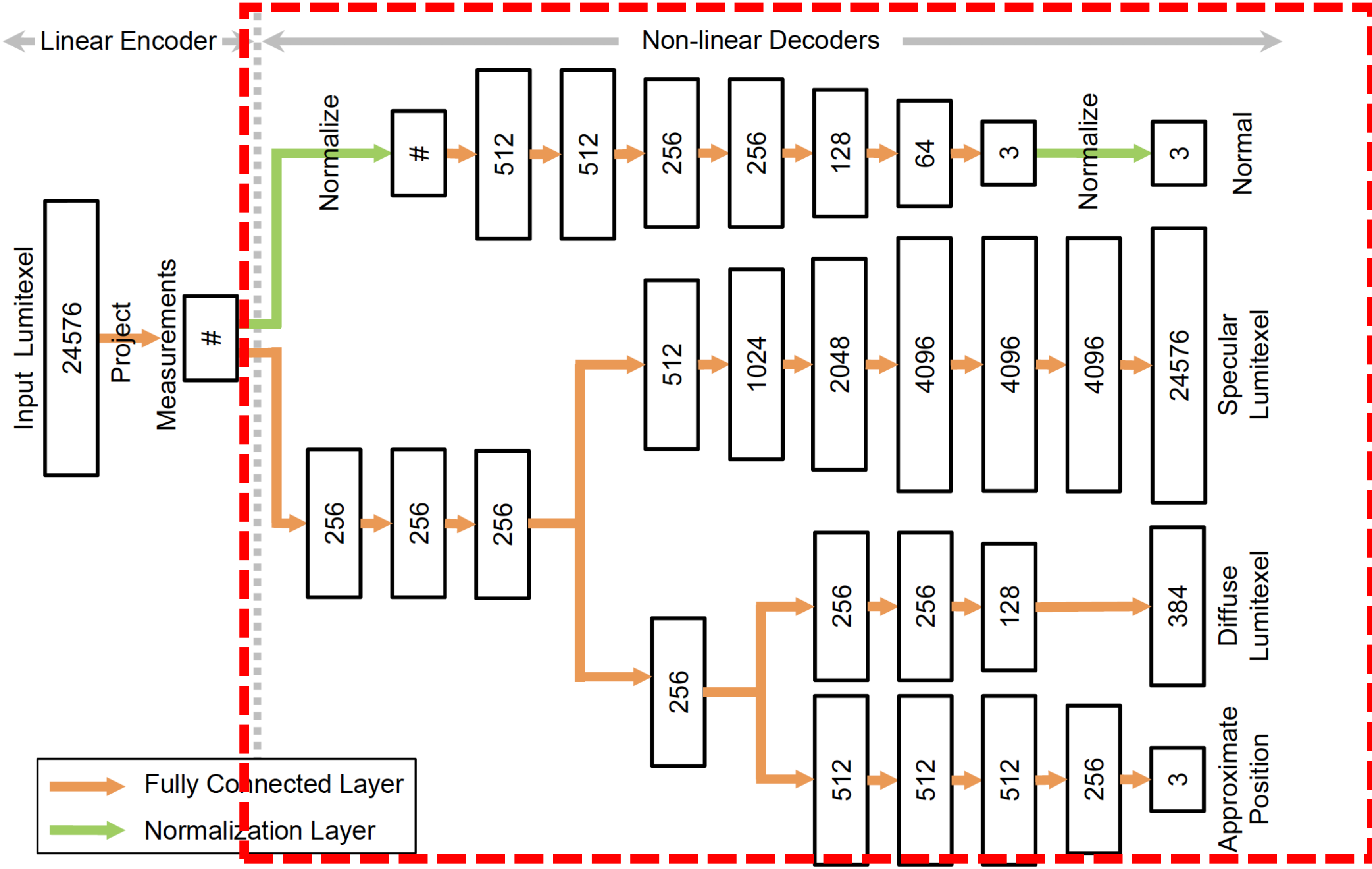
Our Network

- 1 Encoder
  - Physical Capture
- 4 Decoders
  - Computational Reconstruction
- Asymmetric
- Mixed-Domain
- Per-Pixel



Our Network

- 1 Encoder
  - Physical Capture
- 4 Decoders
  - Computational Reconstruction
- Asymmetric
- Mixed-Domain
- Per-Pixel



Loss Function

$$L = \lambda_d L_d(m_d) + \lambda_s L_s(m_s) + \lambda_n L_n(\mathbf{n}) + \lambda_p L_p(\mathbf{p}).$$

Diffuse Lumitexel $L_d(m_d) = \Sigma_l [m_d(l) - \tilde{m}_d(l)]^2,$

Specular Lumitexel $L_s(m_s) = \Sigma_l [\log(1 + m_s(l)) - \log(1 + \tilde{m}_s(l))]^2,$

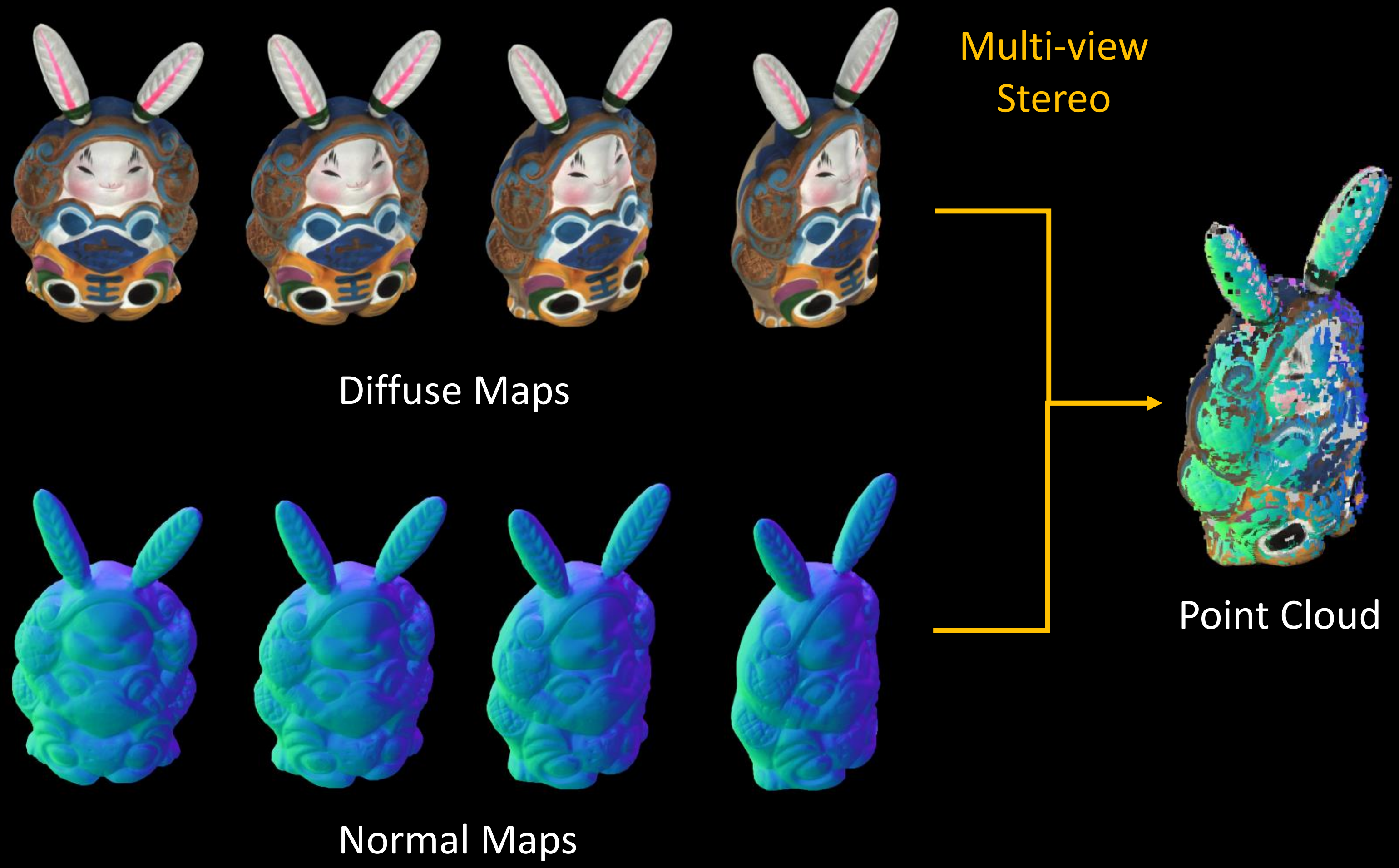
Normal $L_n(\mathbf{n}) = ||\mathbf{n} - \tilde{\mathbf{n}}||_2,$

Approximate Position $L_p(\mathbf{p}) = ||\mathbf{p} - \tilde{\mathbf{p}}||_2,$

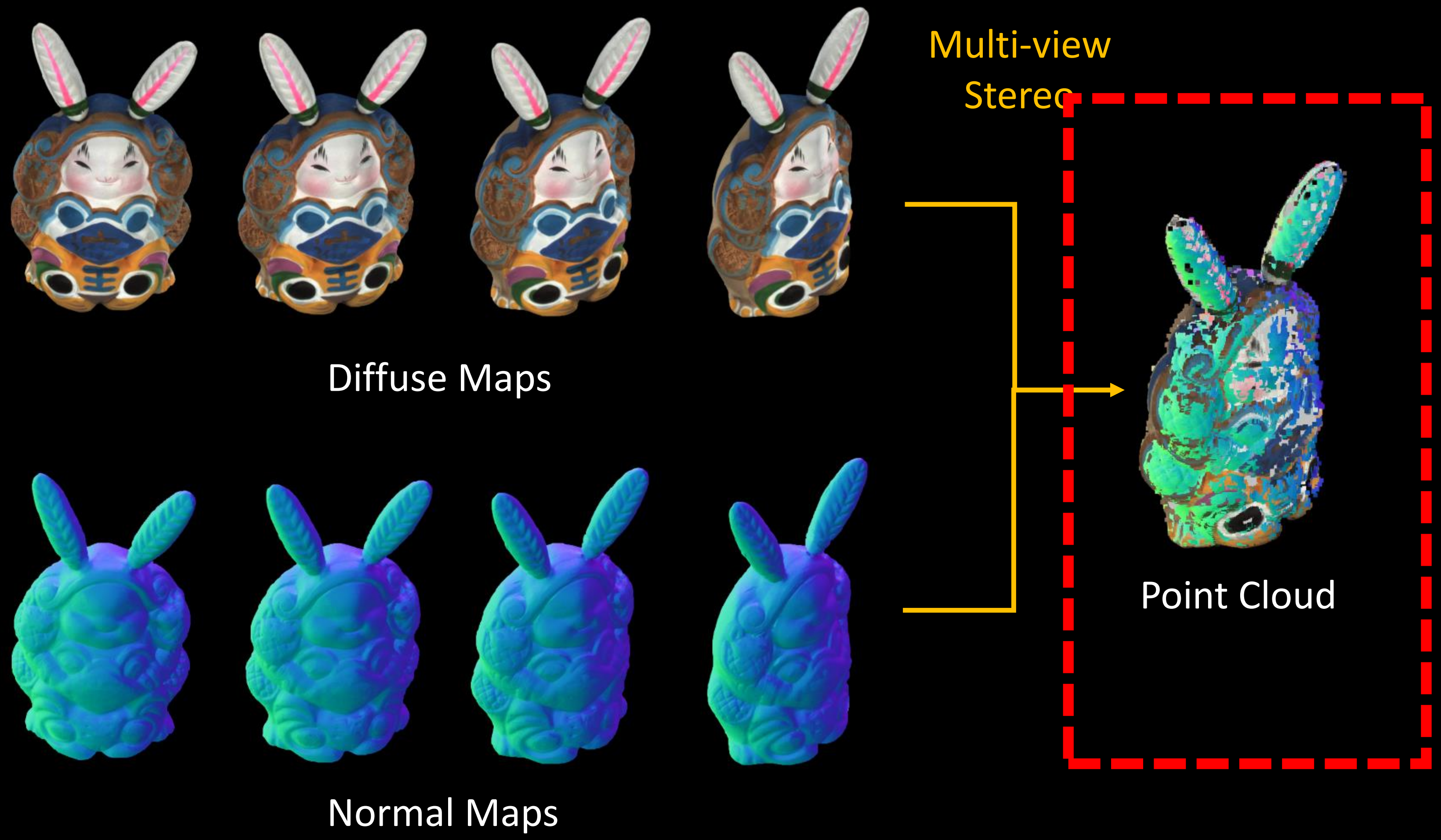
Training

- 200 Million Synthetic Lumitexels
  - Random Position / Local Frame / BRDF Parameters (Anisotropic GGX)
  - Based on Calibration Data
- To Increase Robustness
  - Add Gaussian Noise to Simulated Measurements
  - 10% Dropout Rate to fc Layers

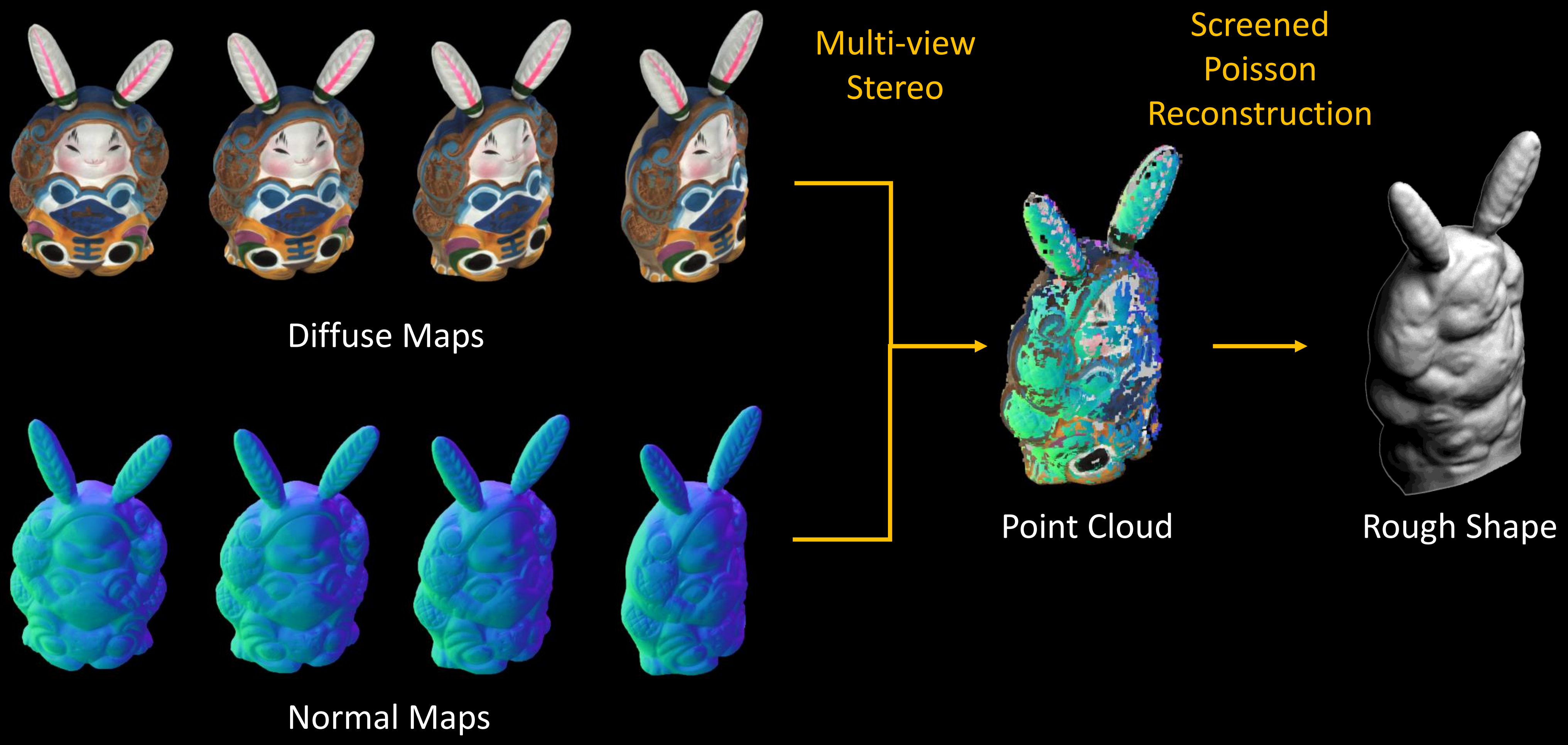
Geometry Reconstruction



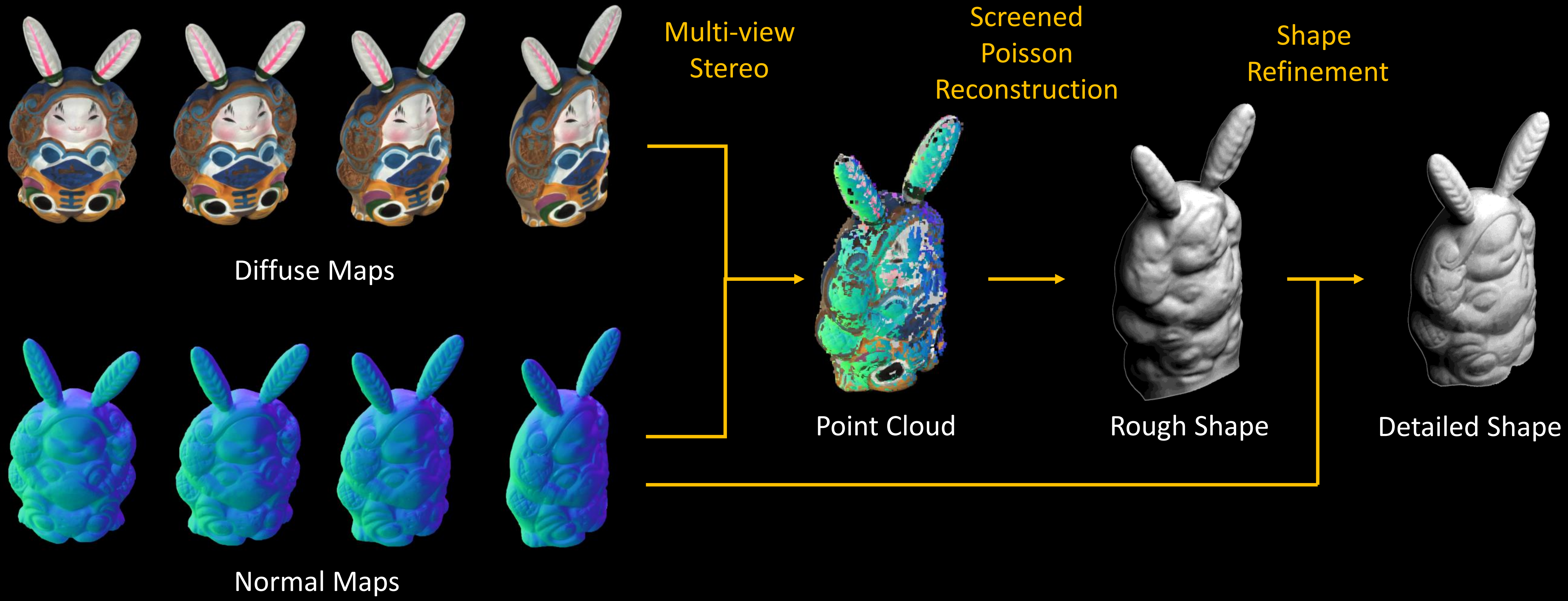
Geometry Reconstruction



Geometry Reconstruction



Geometry Reconstruction



Reflectance Reconstruction

- Input:
  - Decoded Lumitexel
  - 3D Position
- Output:
  - BRDF Parameters (Diffuse / Specular Albedo, Roughnesses, Normal, Tangent)
- Non-linear Optimization using L-BFGS-B

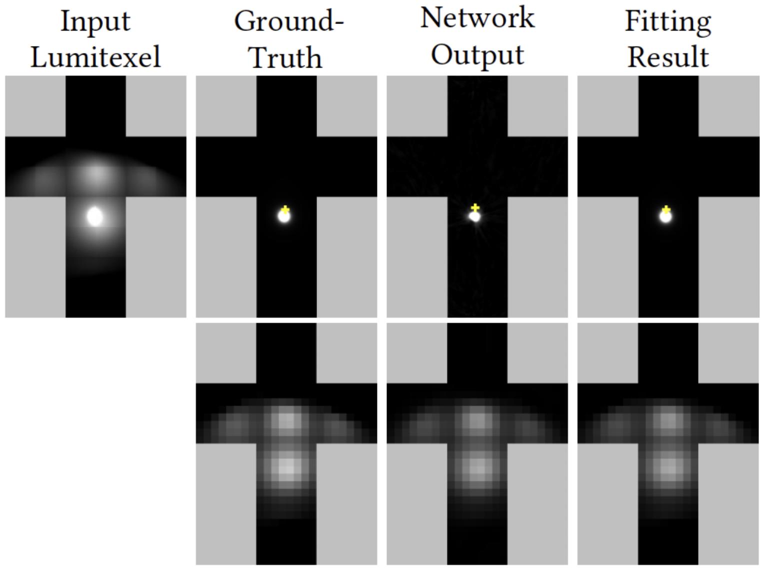
Results

Statistics

• Training:	70 hours
• # Lighting Patterns:	16(isotropic)~32(anisotropic)
• Per-view Acquisition:	7~15 seconds
• Total Acquisition (24 views):	6 minutes
• Decoding:	15 minutes
• Shape Reconstruction:	45 minutes
• Reflectance Fitting:	2 hours

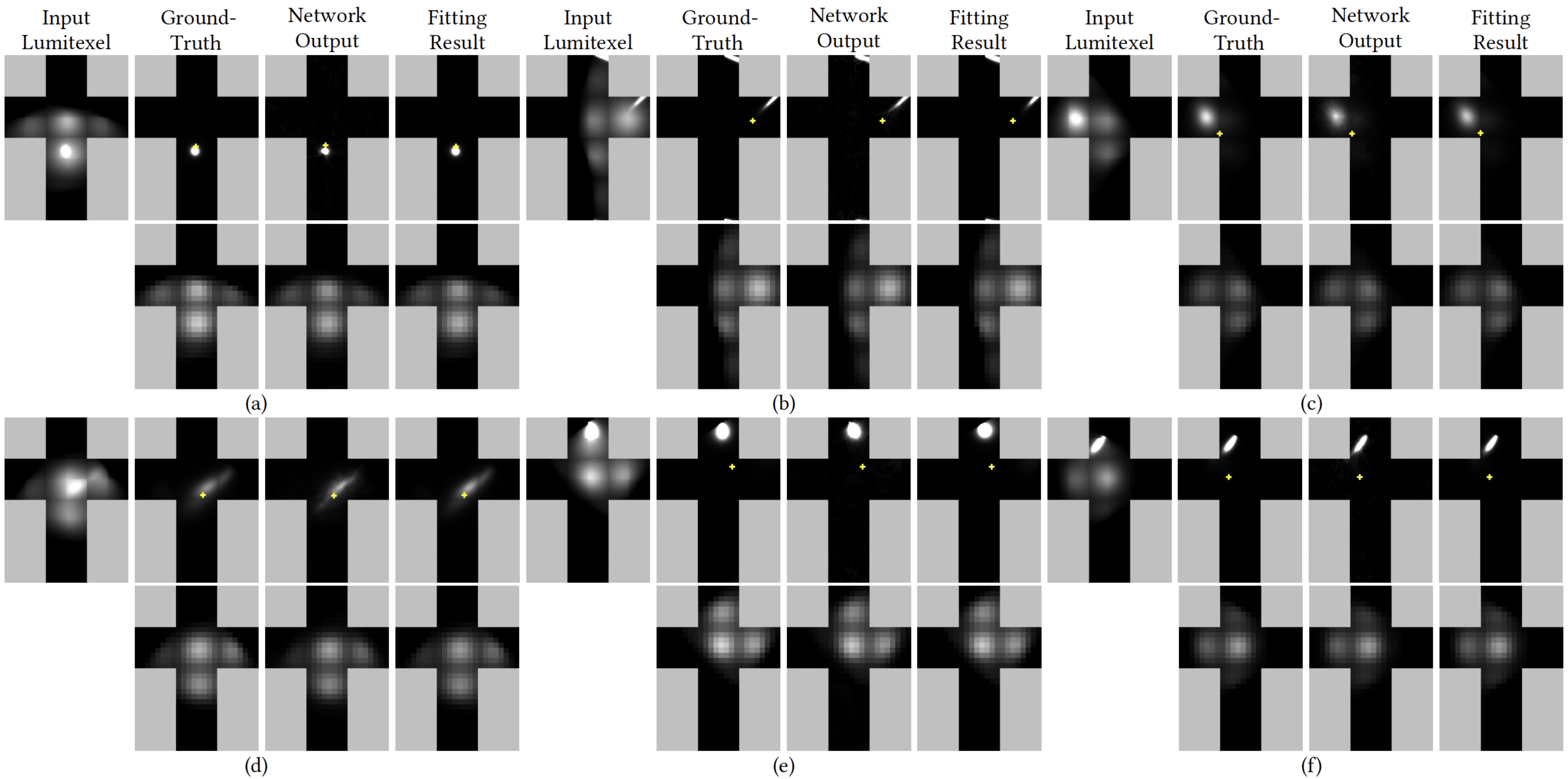
Lighting Patterns

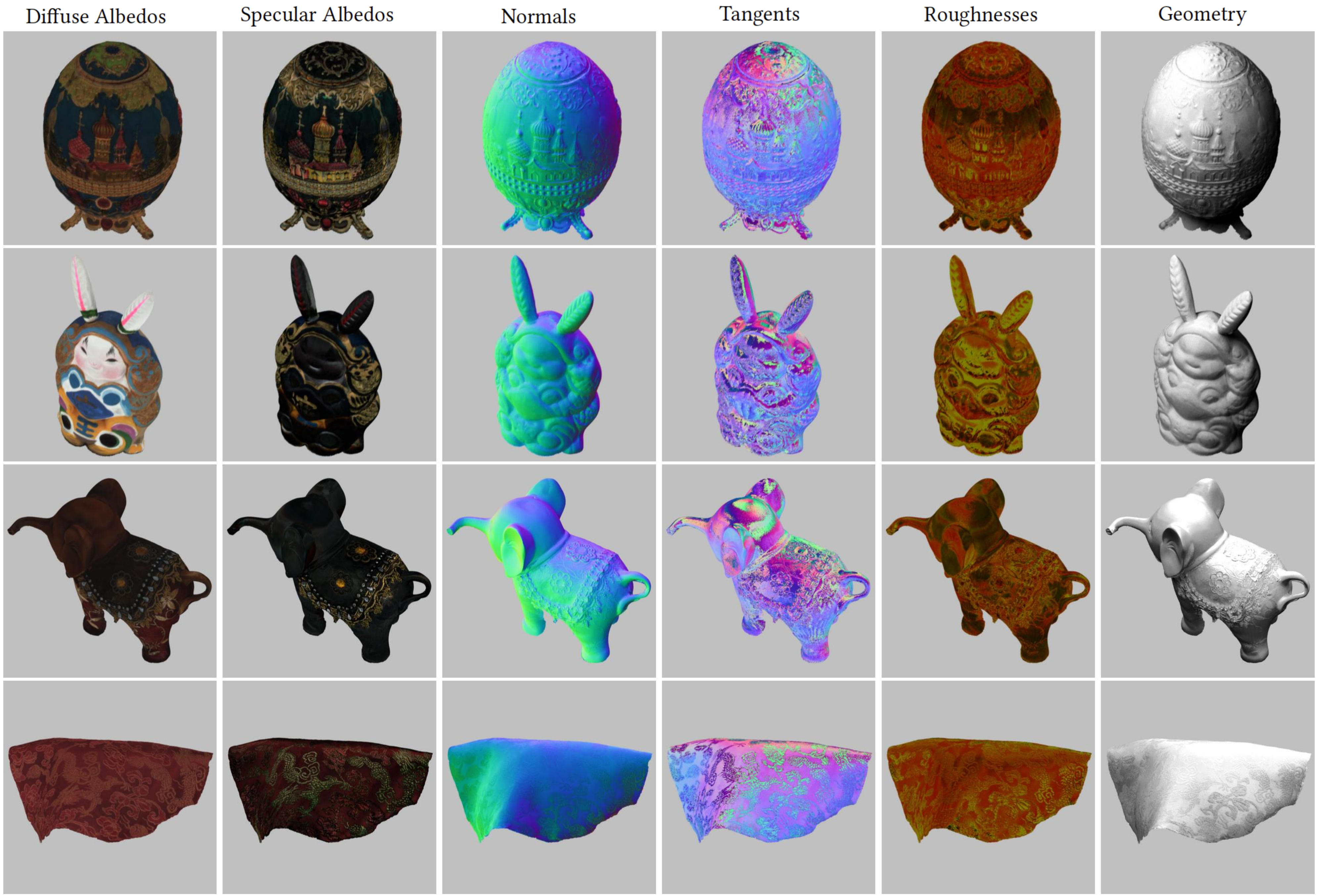
Network Results



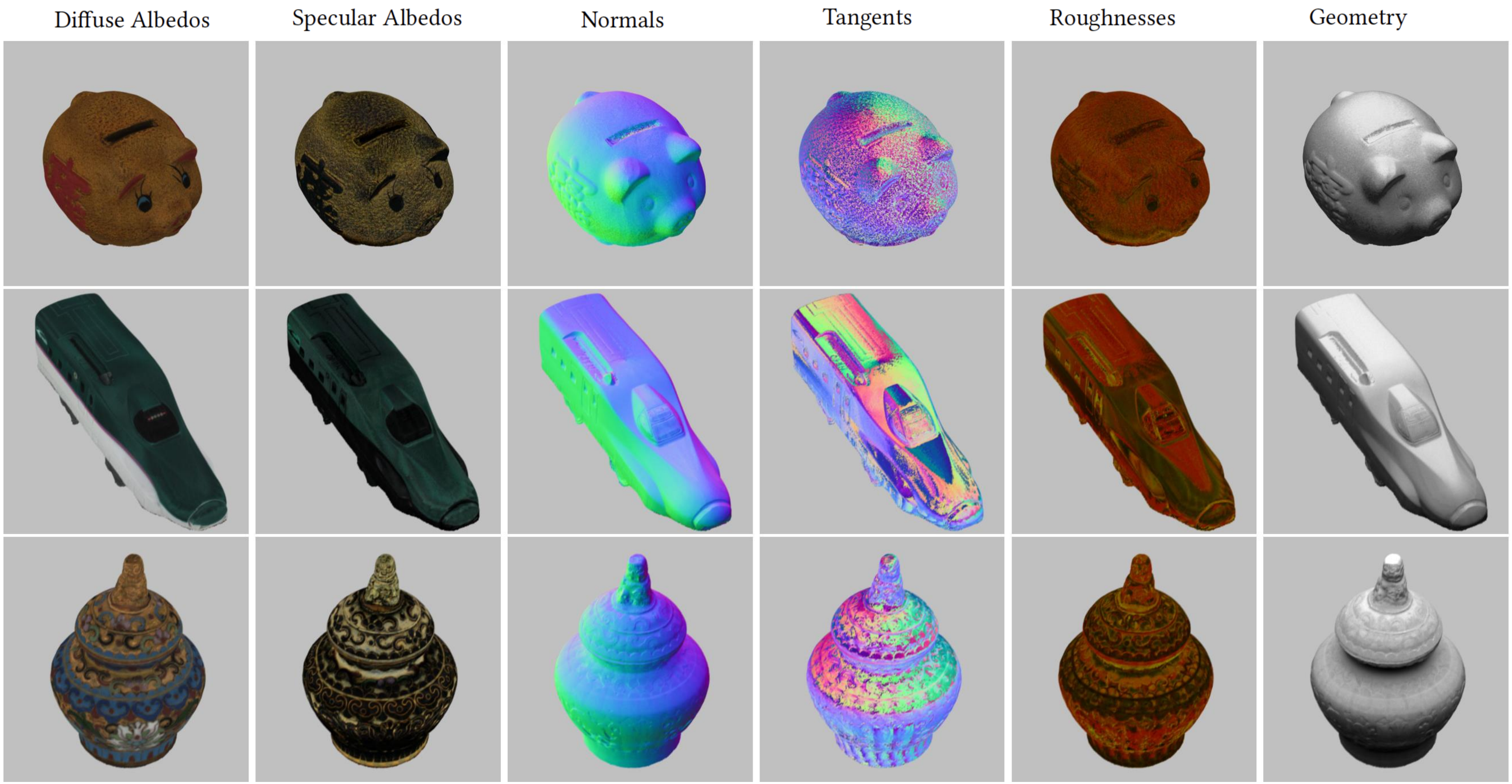
(a)

Network Results

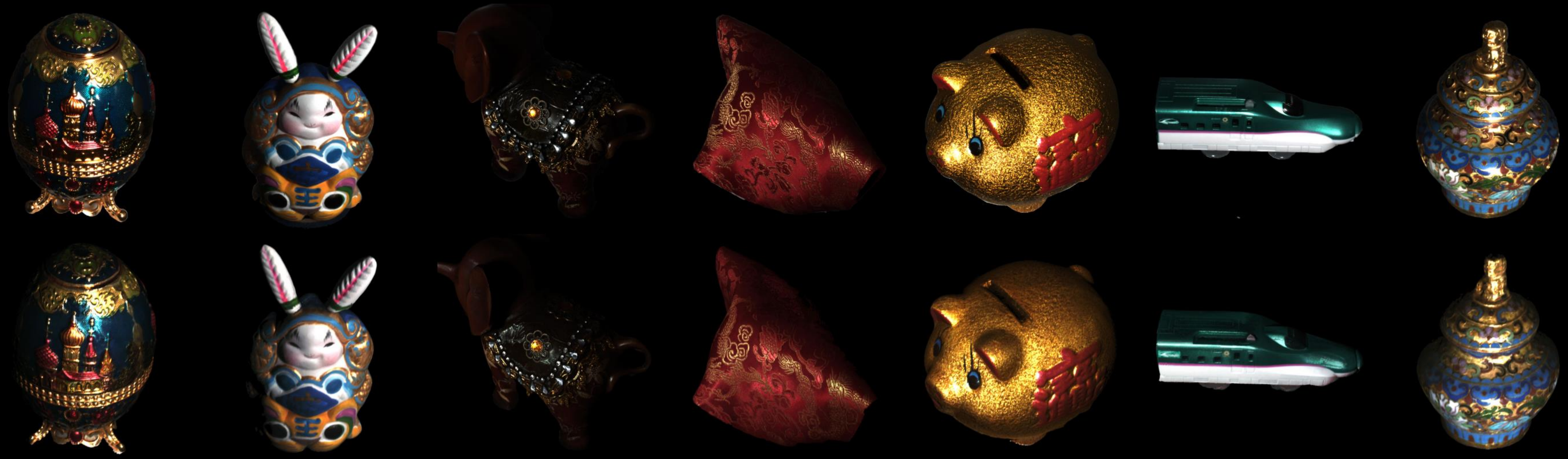




Diffuse Albedos
Specular Albedos
Normals
Tangents
Roughnesses
Geometry


Diffuse Albedos
Specular Albedos
Normals
Tangents
Roughnesses
Geometry

Validation Results



Limitations

- No Explicit Modeling of Inter-reflection / Self-shadowing
- Cannot Recover Appearance Substantially Deviated from Training Samples
- Cannot Reconstruct Details not Observed from Sampled Views

Conclusions & Future Work

- Deep-Learning-Based Framework for Efficient, High-quality Acquisition of Unknown Reflectance & Shape

Conclusions & Future Work

- Deep-Learning-Based Framework for Efficient, High-quality Acquisition of Unknown Reflectance & Shape
- High-quality Photometric Stereo for General Anisotropic Reflectance under Controlled Illumination
  - Average Normal Prediction Error 3.8°

Conclusions & Future Work

- Deep-Learning-Based Framework for Efficient, High-quality Acquisition of Unknown Reflectance & Shape
- High-quality Photometric Stereo for General Anisotropic Reflectance under Controlled Illumination

- Inspire More Research on Differentiable Acquisition
- Apply to Existing / Novel Setups
- Exploit View Coherence
- Handle Other Types of Appearance

Acknowledgements

- Anonymous Reviewers
- Yue Dong(MSRA), Xiaohe Ma, Lijian Ge, Jingke Wang, Tong Yang(ZJU)
- Design Connected EOOD (www.designconnected.com)

- National Key Research & Development Program of China (2018YFB1004300)
- National Science Foundation of China (61772457 & U1609215)

Thank you / Merci / Gracias / 謝謝

- Email: cocoa_kang@zju.edu.cn

- Project Website:

Design Considerations

- Approximate Positions
  - Sufficient for Diffuse Albedo Computation
  - Insufficient for Geometry Reconstruction
- Per-Pixel Normal Prediction v.s. Fitting
  - Breaks the Mutual Dependency of Reflectance and Shape Reconstruction
- Lumitexel Prediction v.s. BRDF Parameter Regression
- No Spatial Aggregation in Our Network
  - Exploit State-of-the-Art Related Work
  - Avoid Combinatorial Explosion in Training Data