Microsoft Research 微软亚洲研究院

Data Driven Graphics
What Can We Get from Sparse and Dense Data?

Xin Tong,

Internet Graphics Group, MSRA

2019.5

Outline

- An overview of data driven graphics
  - Set of techniques for different applications
- Key challenges in data driven graphics and our exploration
  - The underline logic/connection behind researches
- Future directions
  - It is your turn…

Outline

- An overview of data driven graphics
- Key challenges in data driven graphics and our exploration
- Future directions

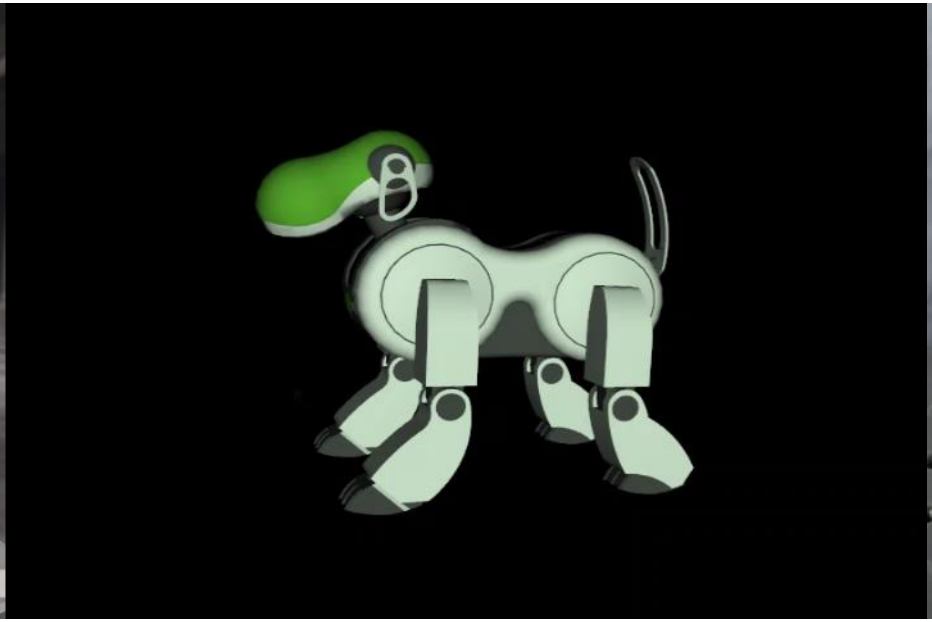
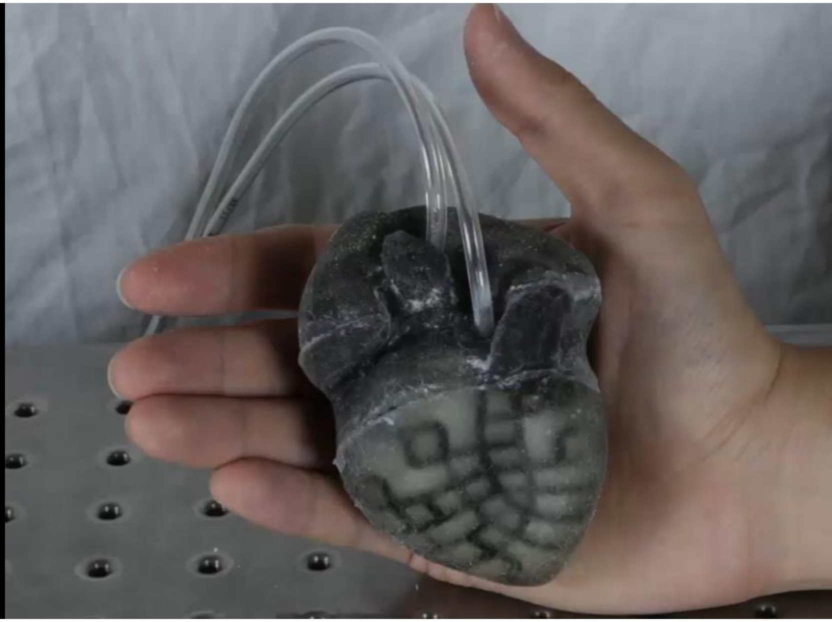
Computer Graphics

- Creating realistic 3D contents in the computer
  - Geometry
  - Materials
  - Lighting effects
  - Dynamics




Physically Based Approach

- Modeling the virtual world by following the geometric & physical rules of the real world

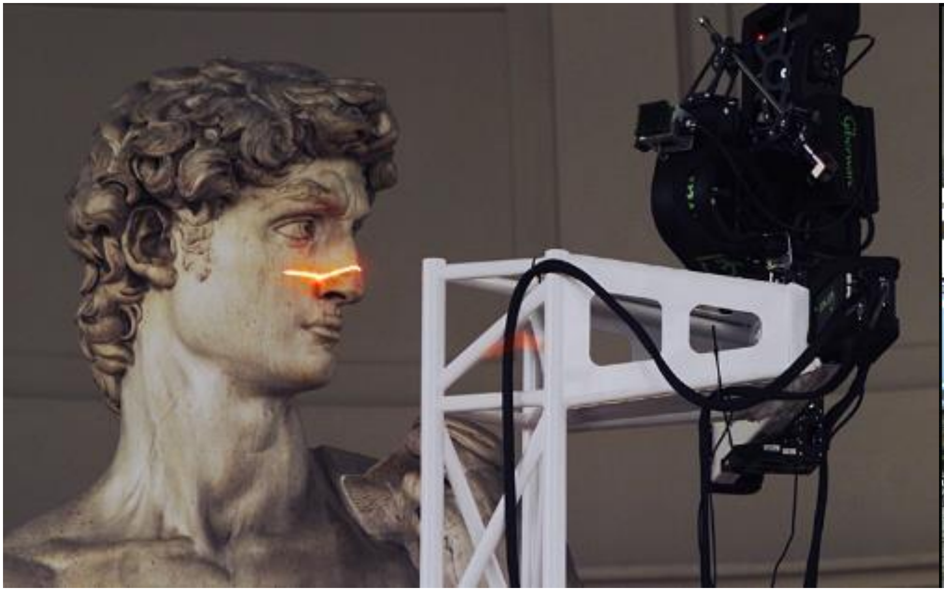
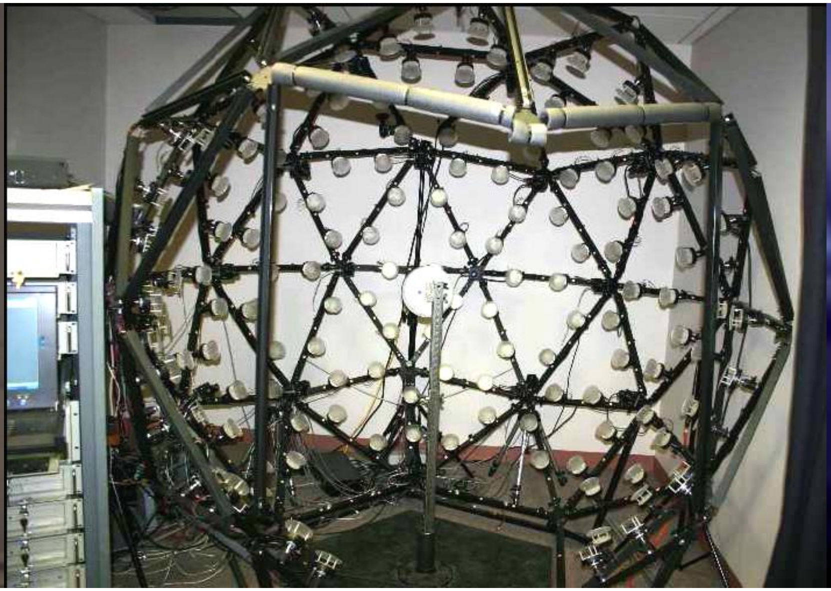
Physically Based Approach

- Modeling the virtual world by following the geometric & physical rules of the real world
  - ☺Compact and clean

  - ☹Computational expensive
  - ☹Huge efforts for modeling the rich details

Data Based Approach

- Densely sampling the target (geometry, material, lighting…) space and reconstructing the results by interpolation


Left: http://illumin.usc.edu/46/michelangelo39s-motion-picture/
Middle: http://www.cs.columbia.edu/CAVE/projects/time_var/time_var.php
Right: http://www.btlnews.com/crafts/visual-fx/vicon-launches-new-facial-motion-capture-system/

Microsoft Research 微软亚洲研究院

Data Based Approach

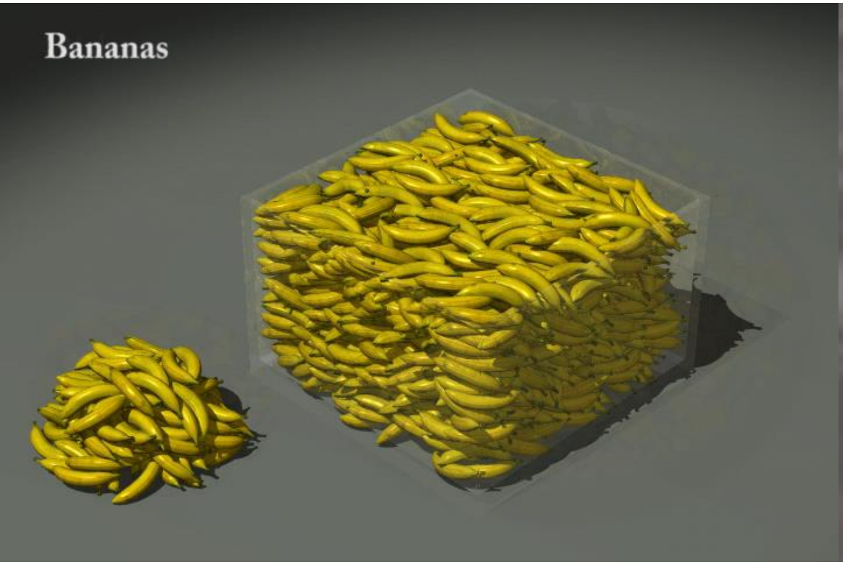
- Densely sampling the target (geometry, material, lighting…) space and reconstructing the results by interpolation
  - ☺Directly capture the data from the real world
  - ☺Fast computation for reconstruction
  - ☺High fidelity results with all details

  - ☹Expensive capturing devices and setup
  - ☹Hugh amount of data
  - ☹Difficult to manipulate and edit

Microsoft Research 微软亚洲研究院

Data Driven Approach

- Inferring the results from an efficient target space model (geometry, materials, lighting…) learned from the data samples



Microsoft Research 微软亚洲研究院

Data Driven Approach

- Inferring the results from an efficient target space model (geometry, materials, lighting…) learned from the data samples
  - ☺High fidelity results
  - ☺Easy to edit and manipulate

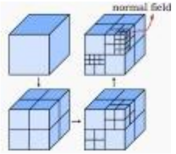
  - ☹How to learn the model of the target space?

Data Driven Approach: Our Efforts

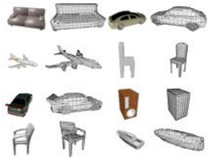
Geometry Modeling

Discrete Element Texture [TOG2011]

Mesh Denoising [TOG2016]



O-CNN [TOG2017]



AO-CNN [TOG2018]

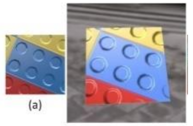
Appearance Modeling

Microfacet Synthesis [TOG2008]

Bootstraping SVBRDF [TOG2010]

Pocket Reflectometry [TOG2011]

Sparse-as-Possible [TOG2016]



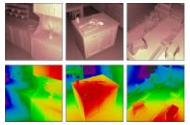
SA-Net [TOG2017]

Rendering

Kernel Nystorm Relighting [TOG2009]

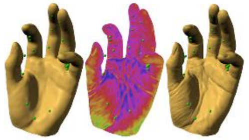
Real Time Global Illumination by Neural Networks [TOG2013]

Religthing by Neural Networks [TOG2015]



DeepTOF [TOG2017]

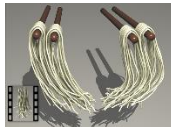
Animation



Detailed Hand Animation [TVCG2012]

Video based Facial Capturing [TOG2014]

Audio-Video Facial Animation [TOG2015]



Dynamic Element Texture [TOG2013]

Microsoft Research 微软亚洲研究院

How to Learn the Model of the Target Space?

Sparse Data

Detailed Hand Animation [TOG2012]

Microfacet Synthesis [TOG2008]

Bootstraping

Mesh Denoising [TOG2016

ctometry [TOG2011]

Video based Facial Ca [TOG2014]

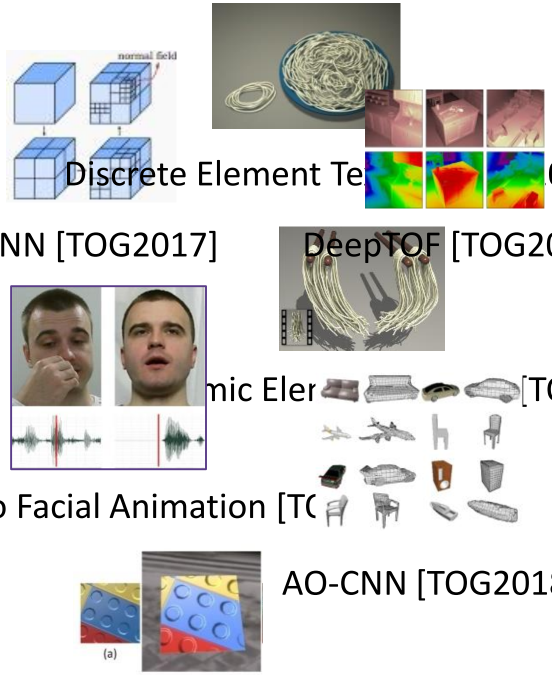
Kernel Nystorm Relighting [TOG2009]

Sparse-as-Possible [TOG2016]

Real Time Global Illumination by Neural Networks [TOG2013]

Religthing by Neural Networks [TOG2015]

Dense Data



Discrete Element Te 011]

O-CNN [TOG2017]

DeepTOF [TOG2017]

mic Eler [TOG2013]

Audio-Video Facial Animation [TC

AO-CNN [TOG2018]

SA-Net [TOG2017]

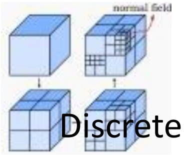
Microsoft Research 微软亚洲研究院

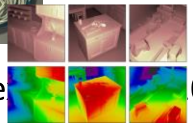
How to Learn a Model of the Target Space?

Sparse Data

Leveraging the priors of the target space for designing compact space model

Dense Data

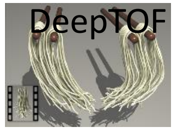


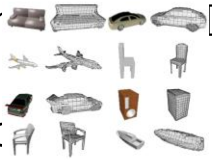


Discrete Element Te... 011]

O-CNN [TOG2017]

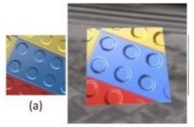
DeepTOF [TOG2017]





...mic Ele... [TOG2013]

Audio-Video Facial Animation [TC...



AO-CNN [TOG2018]

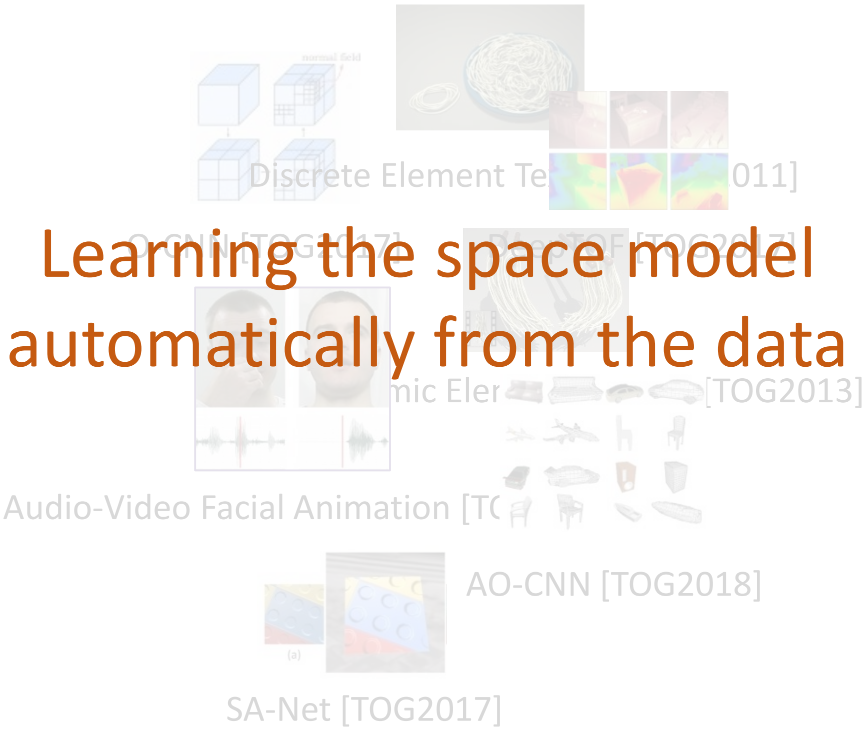
SA-Net [TOG2017]

How to Learn a Model of the Target Space?

Sparse Data

Leveraging the priors of the target space for designing compact space model

Dense Data



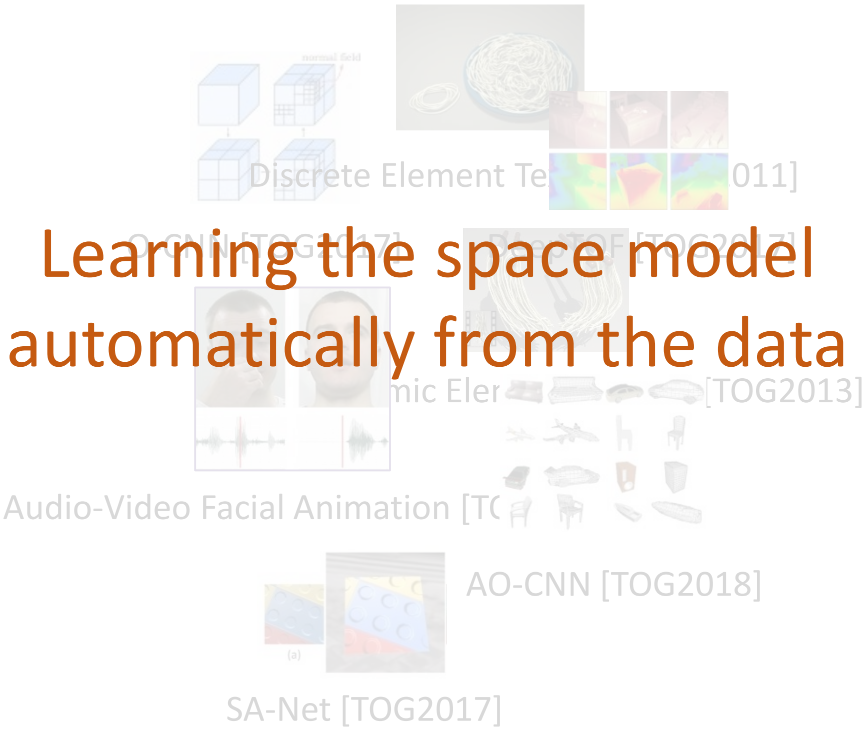
Learning the space model automatically from the data

How to Learn the Model for the Target Space?

Sparse Data

Leveraging the priors of the target space for designing compact space model

Dense Data



Learning the space model automatically from the data

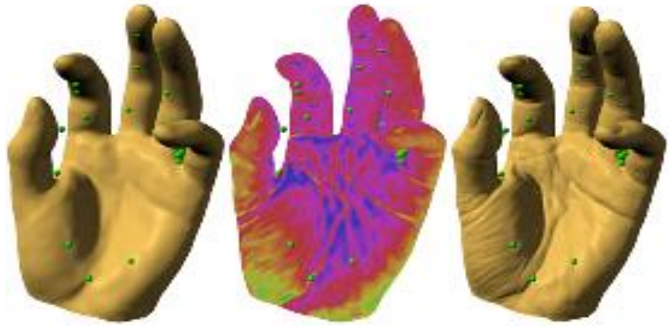
Challenges

- How to design the compact model based on the prior knowledge?

Our Efforts

- How to design the compact model based on the prior knowledge?
- Some strategies: sparse, local, decomposition…

Sparse as Possible SVBRDF Acquisition
[TOG 2016]



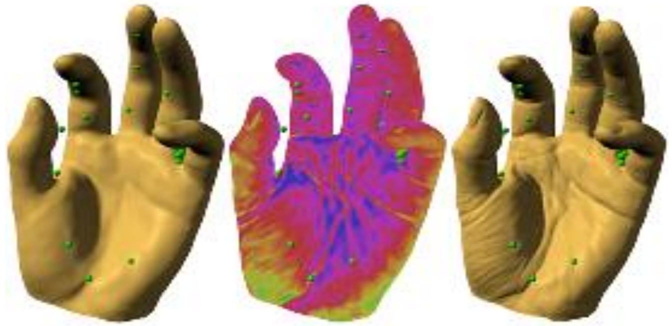
Controllable Hand Deformation from Sparse Examples with Rich Details [SCA 2011]

Microsoft Research 微软亚洲研究院

Our Efforts

- How to design the compact model based on the prior knowledge?
- Some strategies: sparse, local, decomposition…

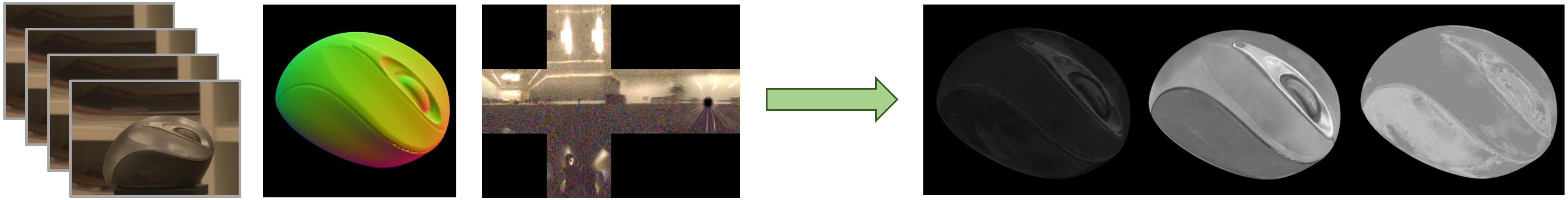
Sparse as Possible SVBRDF Acquisition
[TOG 2016]



Controllable Hand Deformation from Sparse Examples with Rich Details [SCA 2011]

Our Goal

- Capturing high quality SVBRDF from as few as possible images
  - 3D shape is known
  - Lighting is known
  - How many images are needed for reconstructing a SVBRDF?

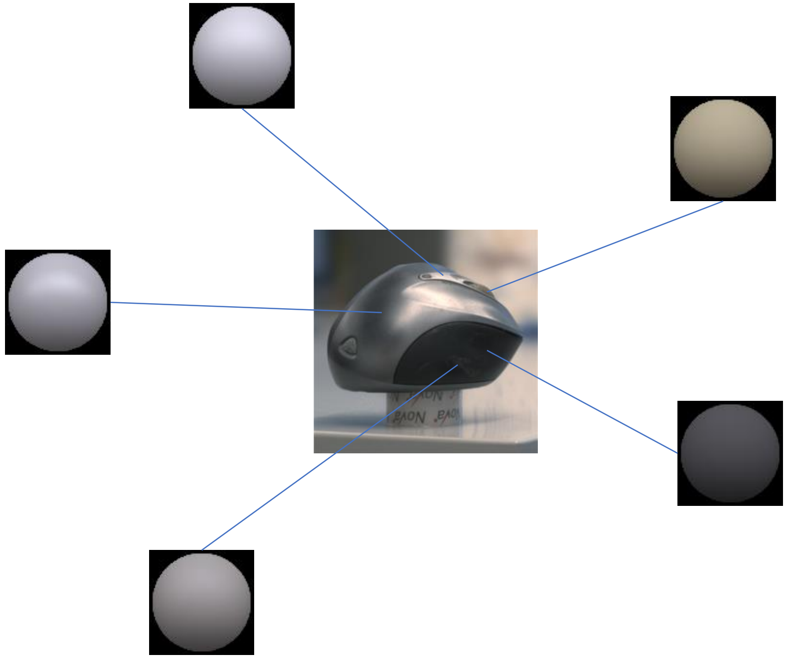


Sparse images | Geometry | Lighting | Surface reflectance (SVBRDF)

Zhiming Zhou, Guojun Chen, Yue Dong, David Wipf, Yong Yu, John Snyder, Xin Tong, *Sparse as Possible SVBRDF Acquisition*, ACM SIGGRAPH ASIA 2016

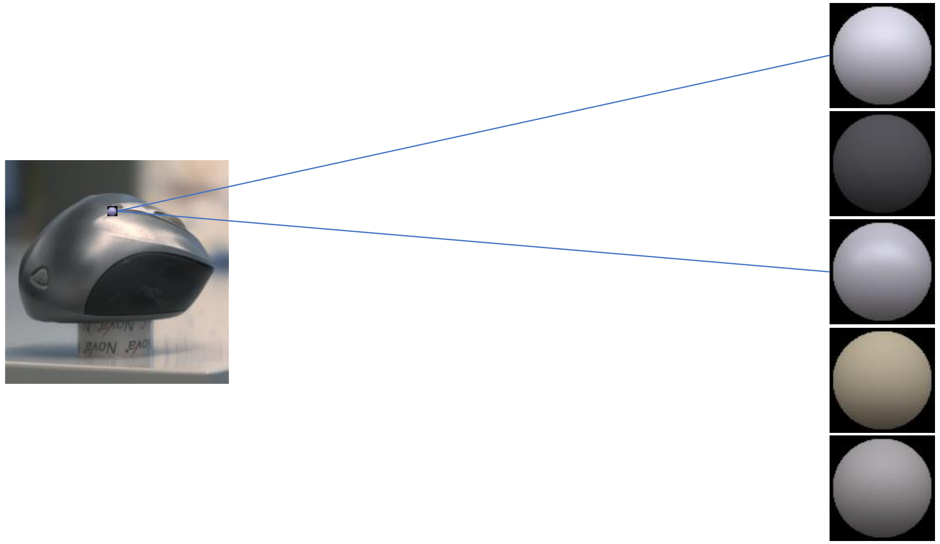
Our Key Observation

- The reflectance of a surface usually formed by **sparse basis materials**



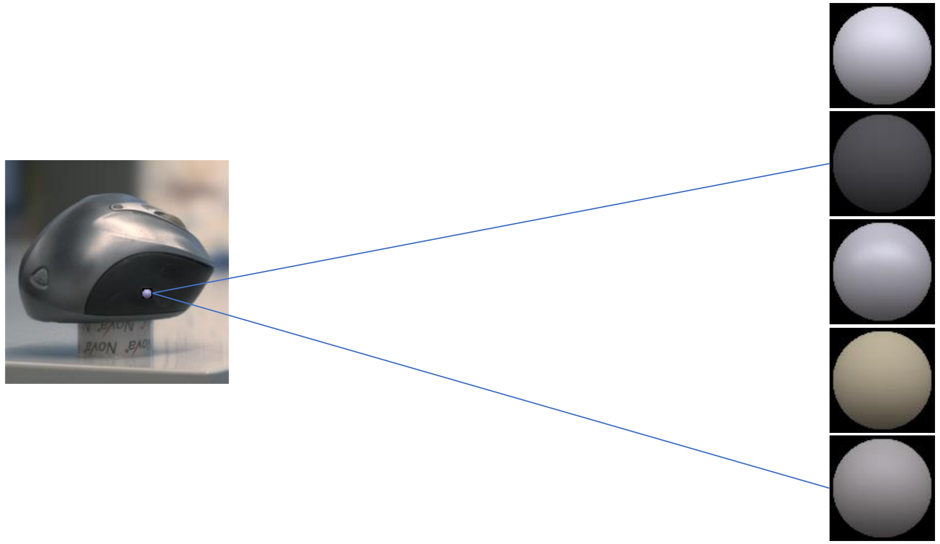
Our Key Observation

- The reflectance of a surface usually formed by **sparse basis materials**
- The BRDF on each point is a **sparse blend** of these basis



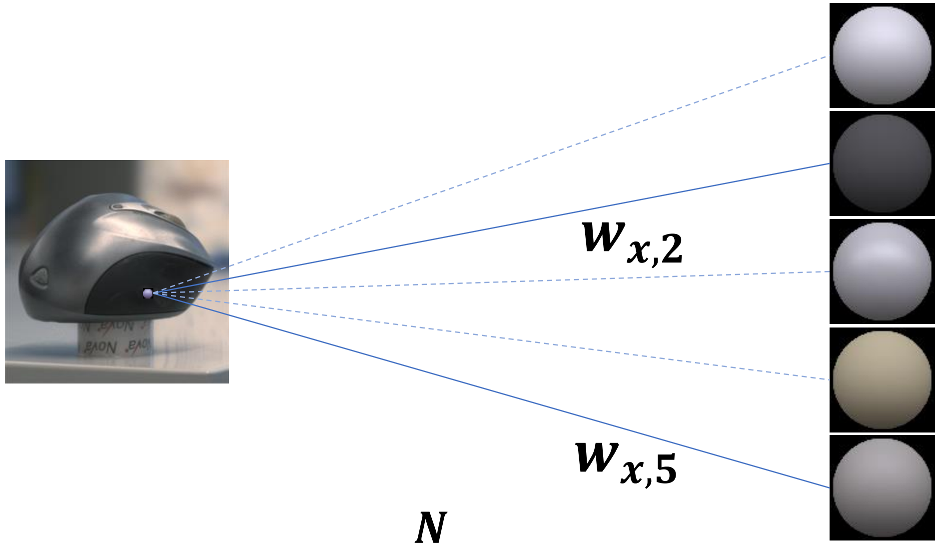
Our Key Observation

- The reflectance of a surface usually formed by **sparse basis materials**
- The BRDF on each point is a **sparse blend** of these basis



Sparse-as-Possible Model

- The reflectance of a surface usually formed by **sparse basis materials**
- The BRDF on each point is a **sparse blend** of these basis

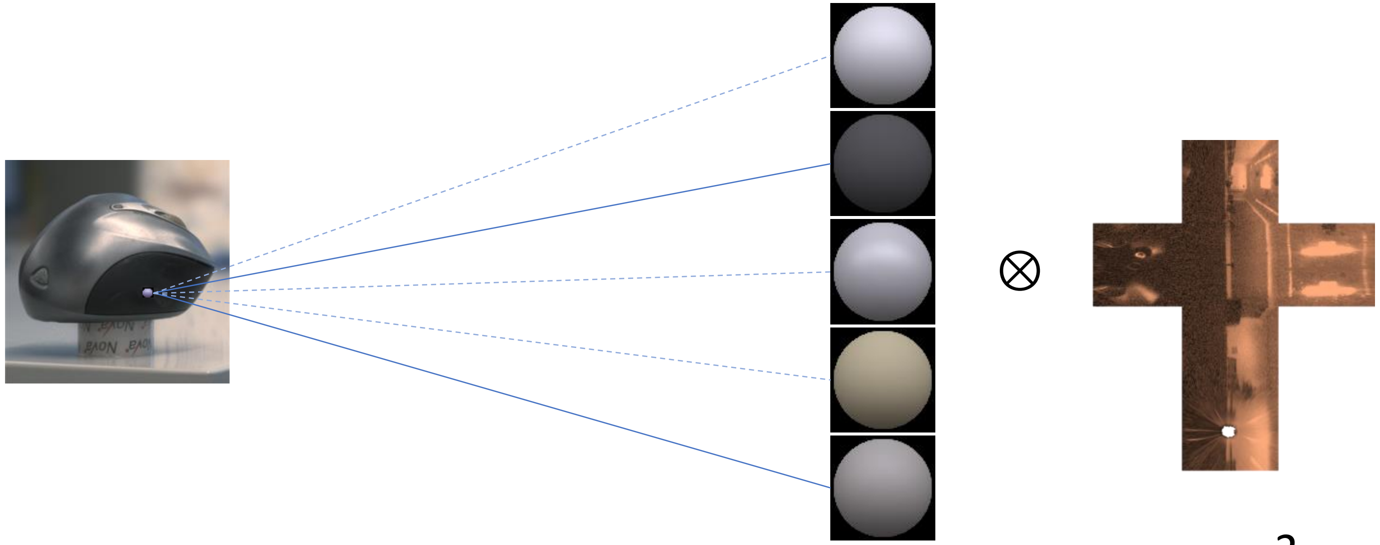


$$B_x = \sum_{i=1}^{N} w_{x,i} * B_i$$

$$\|w_x\|_0 \leq K$$

Technical Challenges

- Solve both sparse blending weights $\boldsymbol{w_{x,i}}$ and basis materials $\boldsymbol{B_i}$

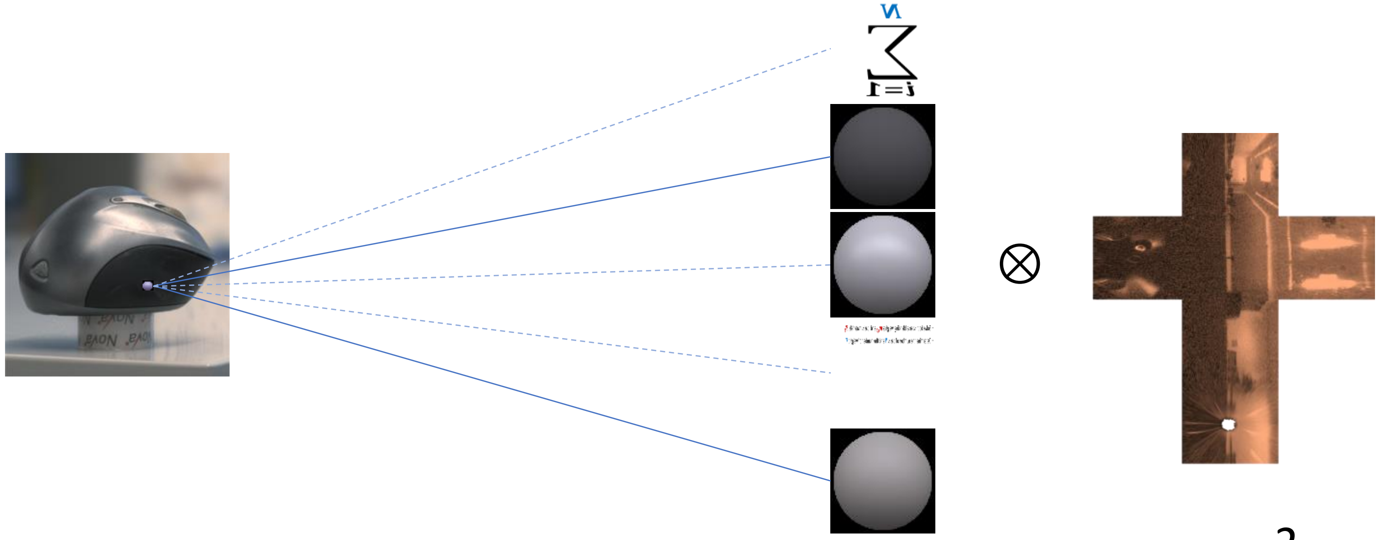


$$argmin \left\| \boldsymbol{S_x} - \left[\sum_{i=1}^{N} \boldsymbol{w_{x,i}} * \boldsymbol{B_i} \right] \otimes \boldsymbol{L} \right\|^2$$

$$\|\boldsymbol{w_x}\|_0 \leq \boldsymbol{K}$$

Technical Challenges

- Solve both sparse blending weights $\boldsymbol{w_{x,i}}$ and basis materials $\boldsymbol{B_i}$
- Determine the number of basis $\boldsymbol{N}$ and the number of weight $\boldsymbol{K}$

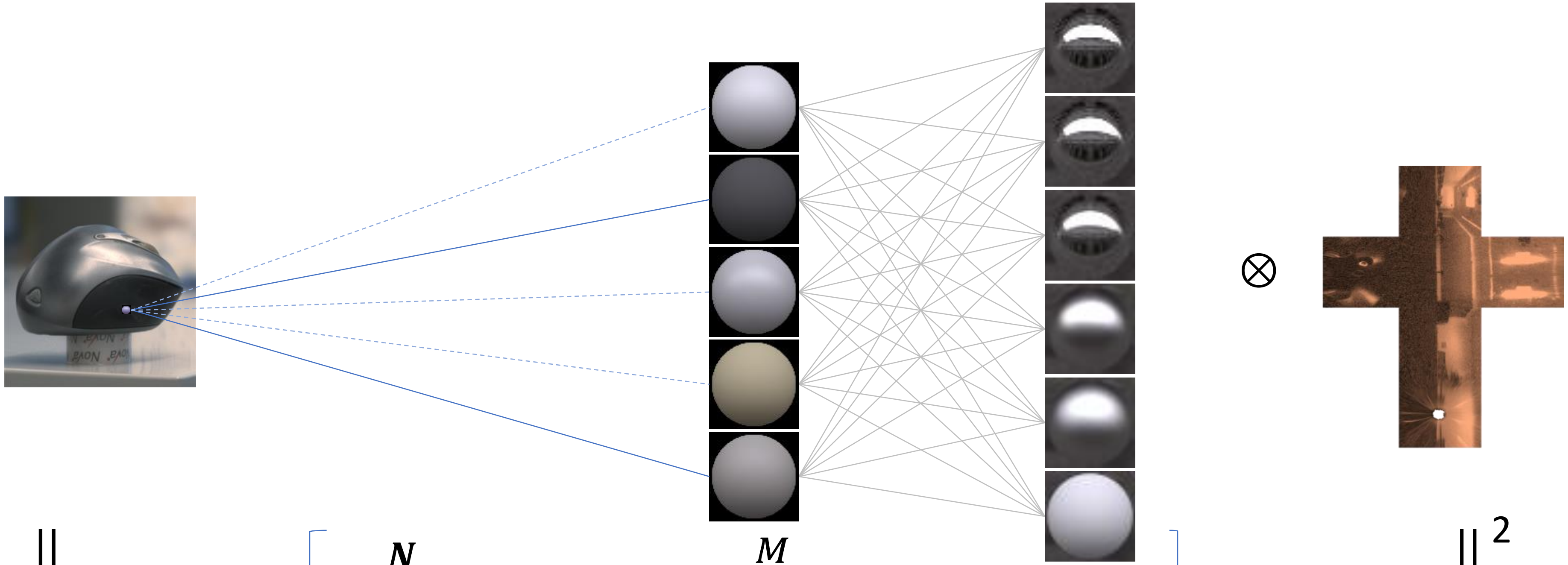


$$\underset{}{argmin} \left\| \boldsymbol{S_x} - \left[\sum_{i=1}^{\boldsymbol{N}} \boldsymbol{w_{x,i}} * \boldsymbol{B_i} \right] \otimes \boldsymbol{L} \right\|^2$$

$$\|\boldsymbol{w_x}\|_0 \leq \boldsymbol{K}$$

Basis and Weight Optimization

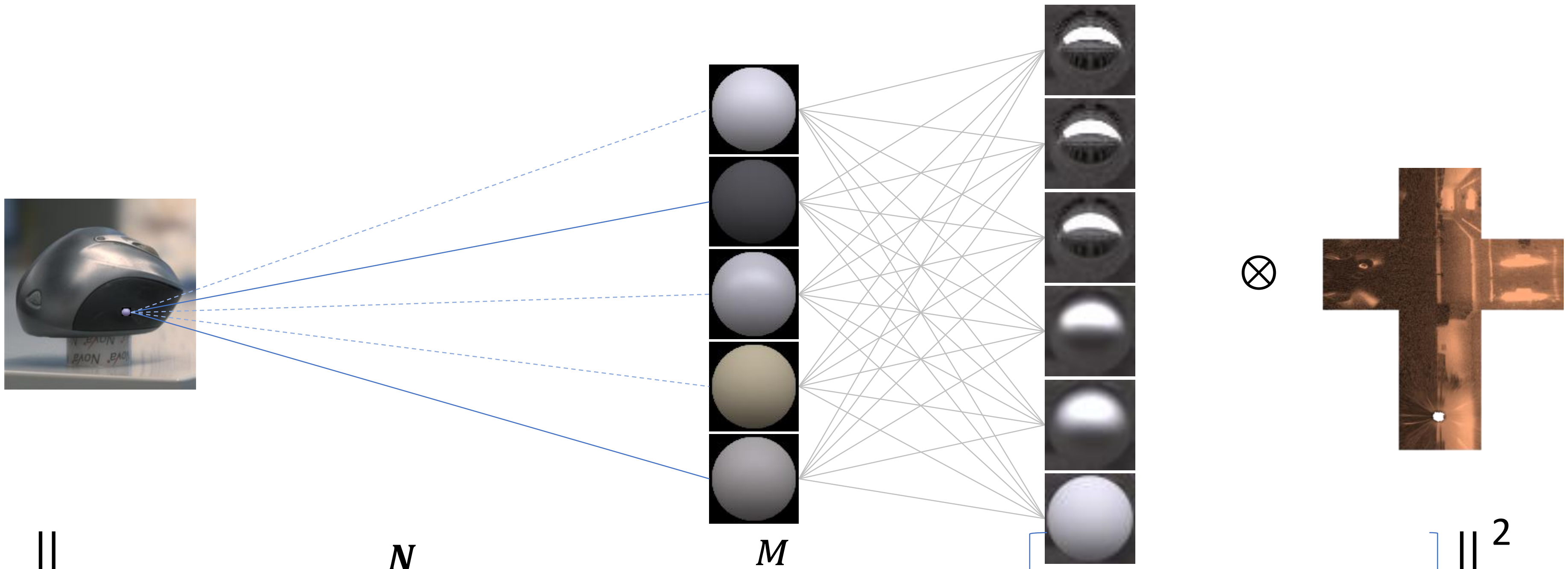
- Model the basis as linear combination of known generic BRDF basis
  - Cook-Torrance BRDFs with different roughness and Fresnel



$$\underset{\|w_x\|_0 \leq K}{argmin} \left\| S_x - \left[\sum_{i=1}^{N} w_{x,i} * \sum_{j=1}^{M} b_{i,j} * G_i \right] \otimes L \right\|^2$$

Basis and Weight Optimization

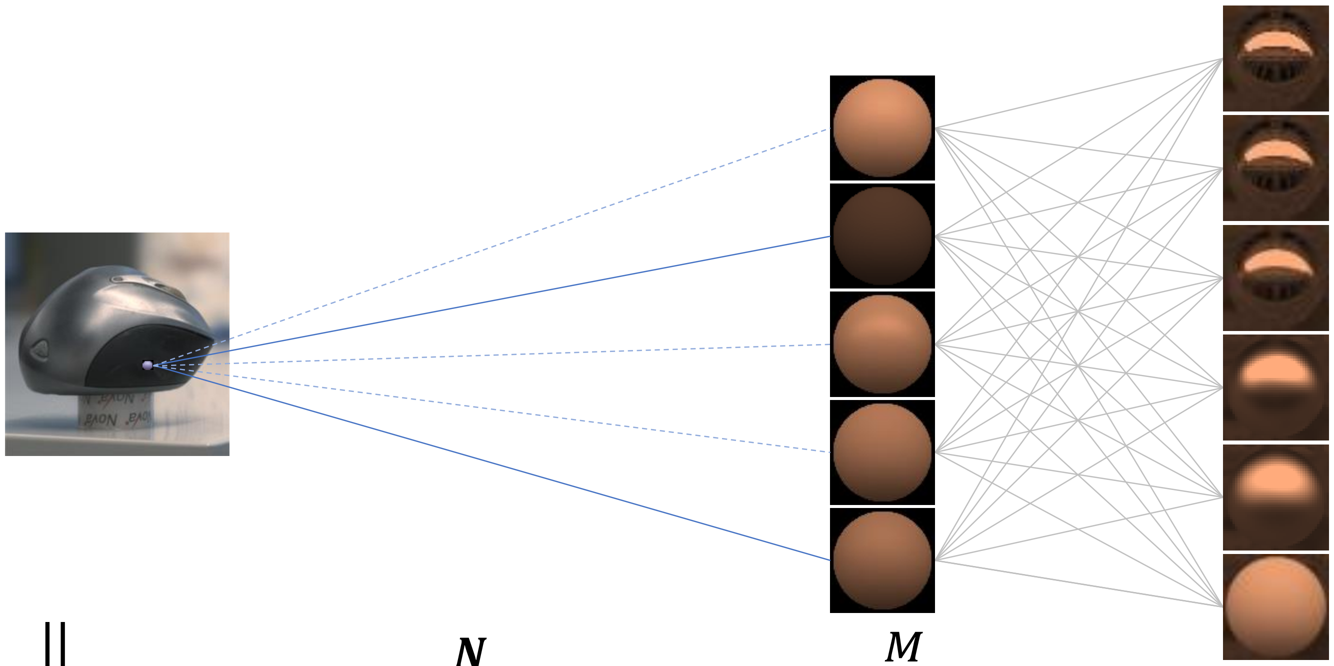
- Rendering generic BRDF basis under given lighting as prediction basis



$$\mathop{argmin}_{\|\boldsymbol{w_x}\|_0 \le \boldsymbol{K}} \left\| \boldsymbol{S_x} - \sum_{i=1}^{N} \boldsymbol{w_{x,i}} * \sum_{j=1}^{M} \boldsymbol{b_{i,j}} * \left[\boldsymbol{G_i} \otimes \boldsymbol{L} \right] \right\|^2$$

Basis and Weight Optimization

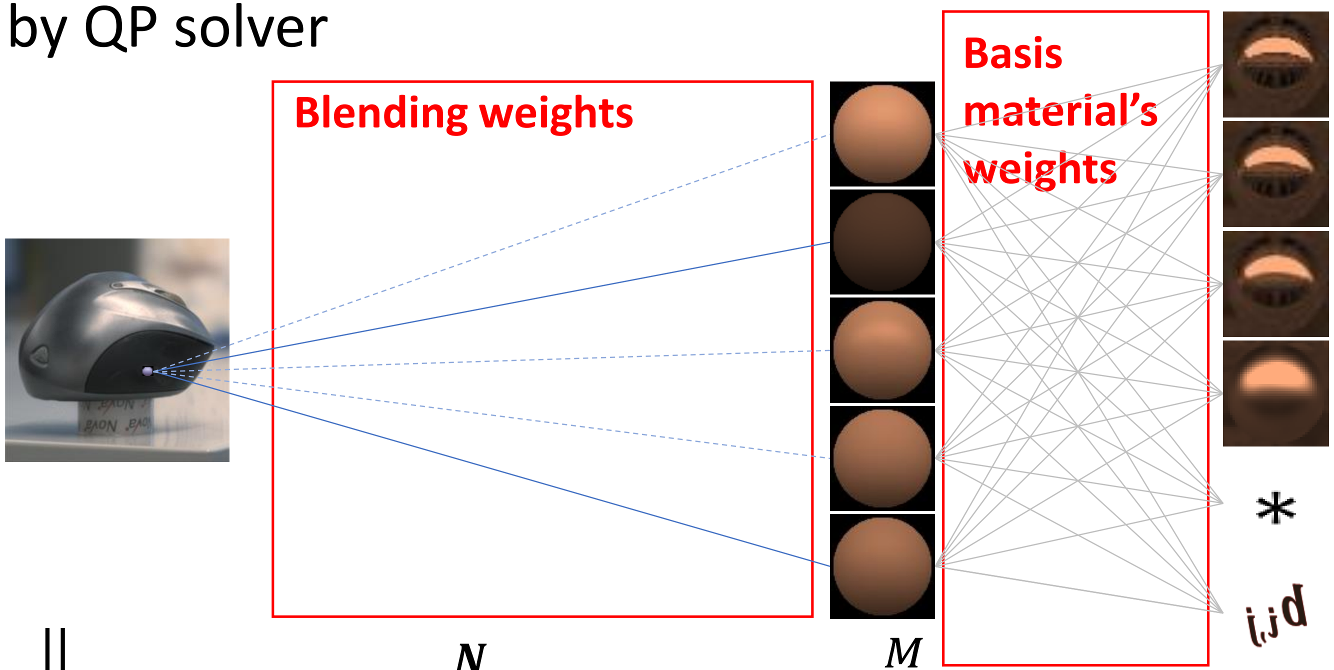
- Rendering generic BRDF basis under given lighting as prediction basis



$$\underset{\|\boldsymbol{w_x}\|_0 \leq \boldsymbol{K}}{argmin} \left\| \boldsymbol{S_x} - \sum_{i=1}^{N} \boldsymbol{w_{x,i}} * \sum_{j=1}^{M} \boldsymbol{b_{i,j}} * \boldsymbol{S_{j_x}} \right\|^2$$

Basis and Weight Optimization

- Iteratively solving basis materials' weights and blending weights
  - Linear system in each step
  - Solve by QP solver



$$\underset{\|\boldsymbol{w}_x\|_0 \le \boldsymbol{K}}{argmin} \left\| \boldsymbol{S}_x - \sum_{i=1}^{N} \boldsymbol{w}_{x,i} * \sum_{j=1}^{M} \boldsymbol{b}_{i,j} * \boldsymbol{S}_{j_x} \right\|^2$$

Determining Number of Basis and Weights

- With two additional L0 constraints for exactly sparse solution

$$\underset{K, N, w_x, w_i^*}{argmin} \left\| S_x - \sum_{i=1}^{N} w_{x,i} \cdot \sum_{j=1}^{M} b_{i,j} \cdot {S_j}_x \right\|^2 + \lambda_1 N + \lambda_0 K$$

$$\|w_x\|_0 \le K$$

Determining Number of Basis and Weights

- Progressively increase the number of weights K and basis N
  - Compute the basis and weights for given K, N
  - Repeat until the total energy starts to increase

$$\underset{K, N, w_x, w_i^*}{argmin} \boxed{\begin{matrix} \left\| S_x - \sum_{i=1}^{N} w_{x,i} \cdot \sum_{j=1}^{M} b_{i,j} \cdot S_{j_x} \right\|^2 \\ \|w_x\|_0 \leq K \end{matrix}} + \lambda_1 N + \lambda_0 K$$

Our Analysis

- N BRDF basis can be reconstructed from measurements of multiple surface points

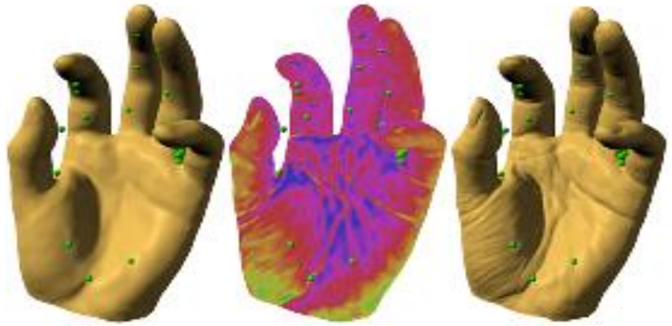
- The number of images needed for reconstructing SVBRDF is always determined by the number of blending weights K!

Real Capture Results: Rendering

Our Efforts

- How to design the compact model based on the prior knowledge?
- Some strategies: sparse, local, decomposition…

Sparse as Possible SVBRDF Acquisition
[TOG 2016]



Controllable Hand Deformation from Sparse Examples with Rich Details [SCA 2011]

Microsoft Research 微软亚洲研究院

Our Goal

- Generating controllable detailed 3D hand animation from sparse 3D pose examples

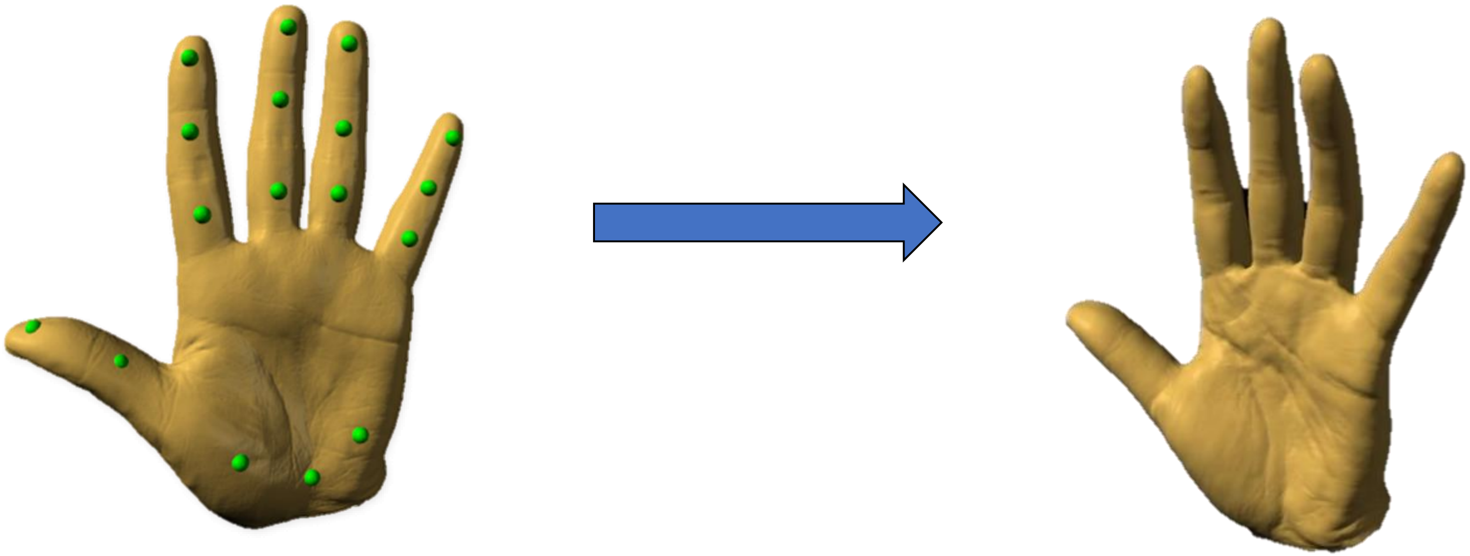
Haoda Huang, Ling Zhao, KangKang Yin, Yue Qi, Yizhou Yu, Xin Tong, *Controllable Hand Deformation from Sparse Examples with Rich Details*, SCA Best Paper, 2011

Key Challenges

- Large DOF of 3D hand motion
  - 21 skeletal degrees of freedom
  - Deformed wrinkle details under different poses
- Very sparse input examples
  - Capturing is difficult

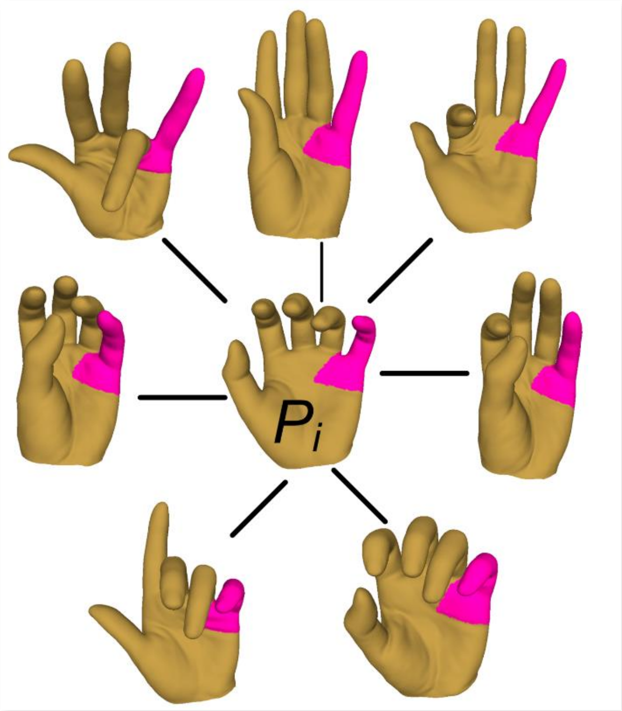
Key Observations

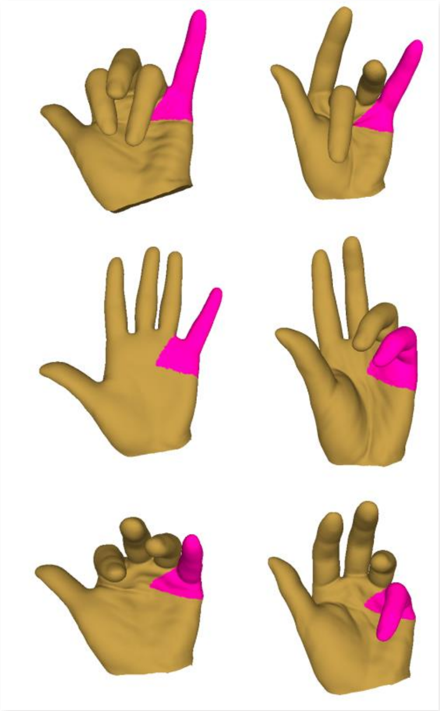
- Leverage the coherence between the motions of different points
  - Transformations of all points can be modeled as functions of control points
  - Can be trained from sparse examples



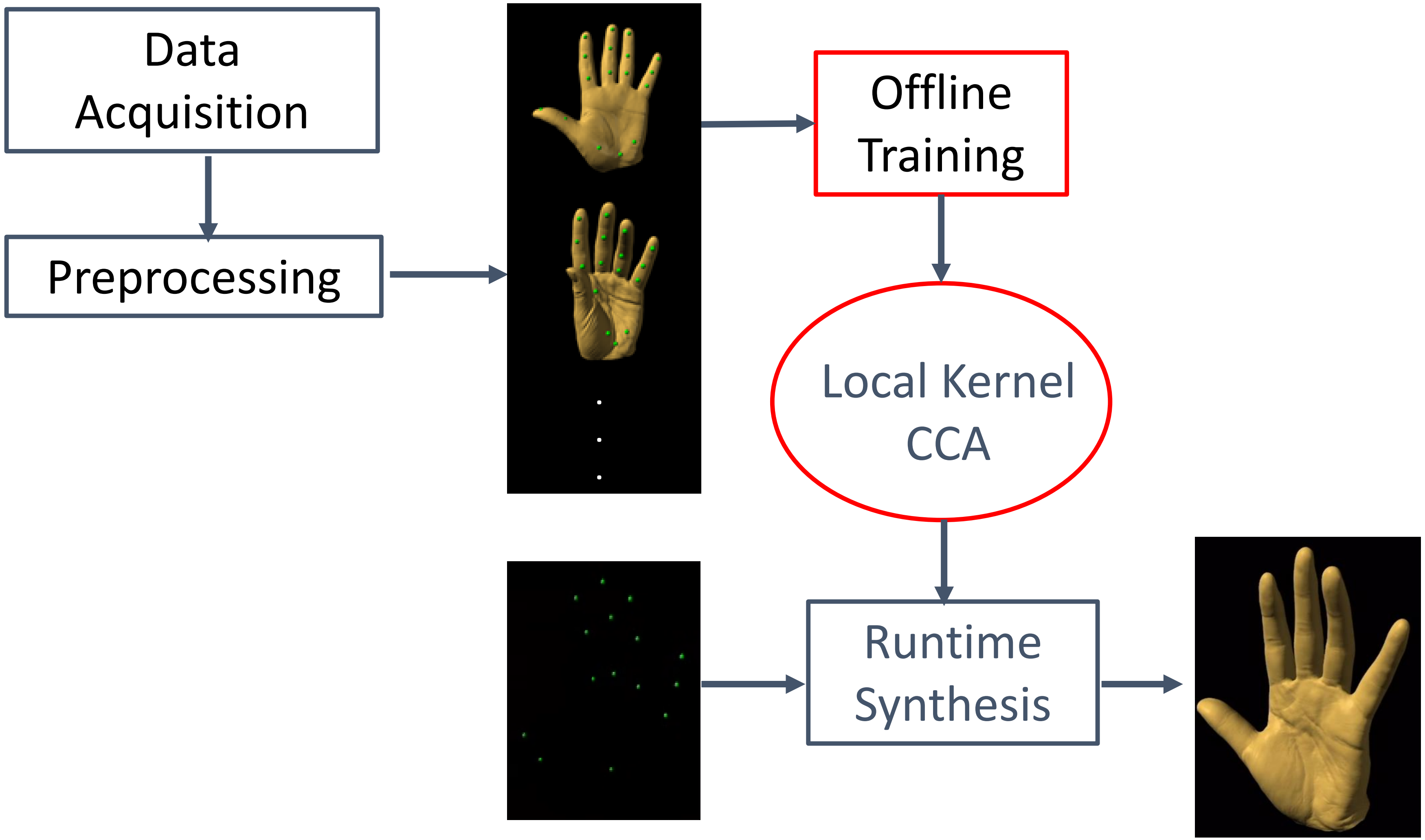
Key Observations

- This function can be modeled by a set of local non-linear functions
  - For local pose space & geometry parts
  - For both coarse level and detail level

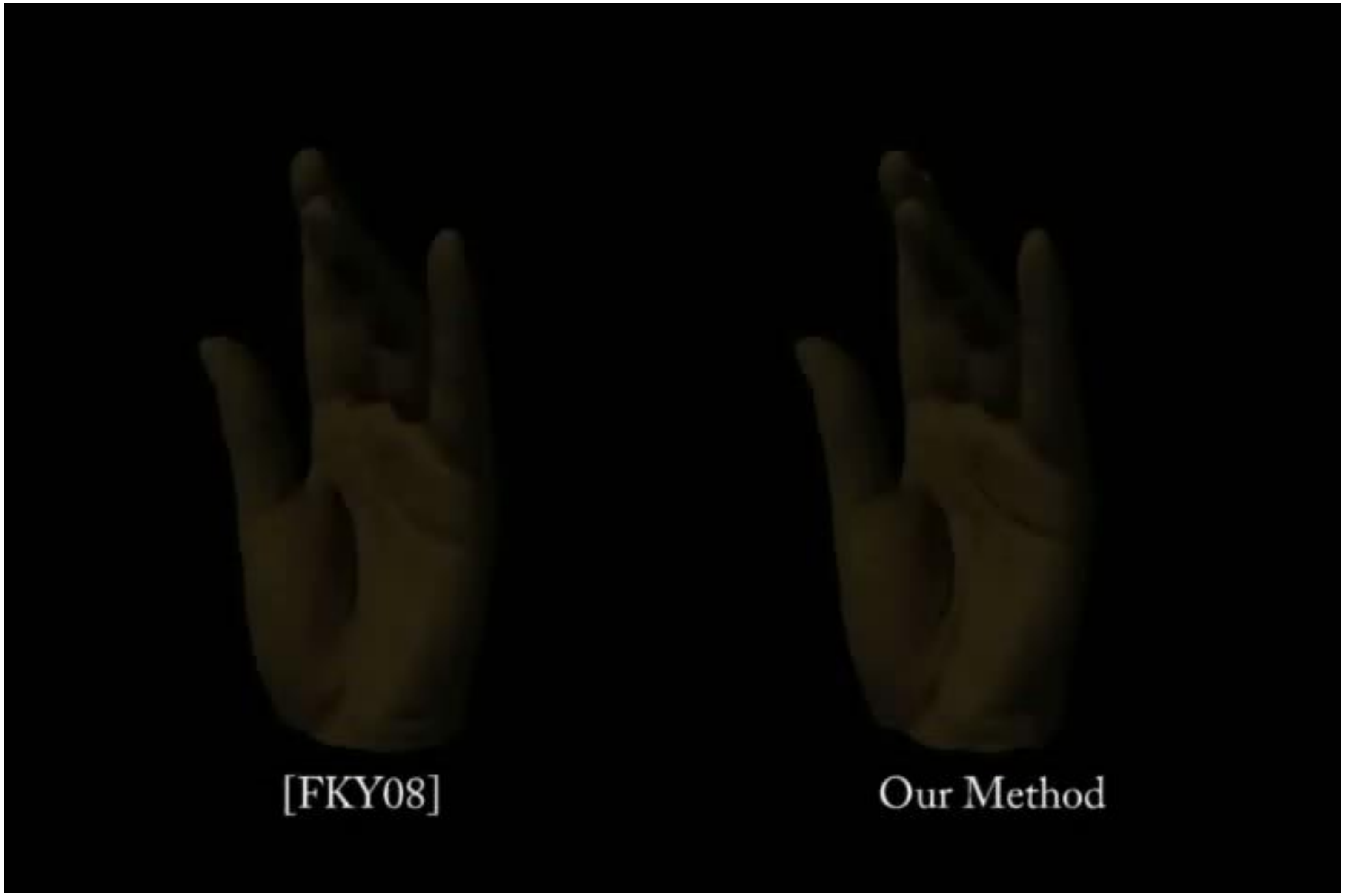




Our Solution



Results: Global vs. Local



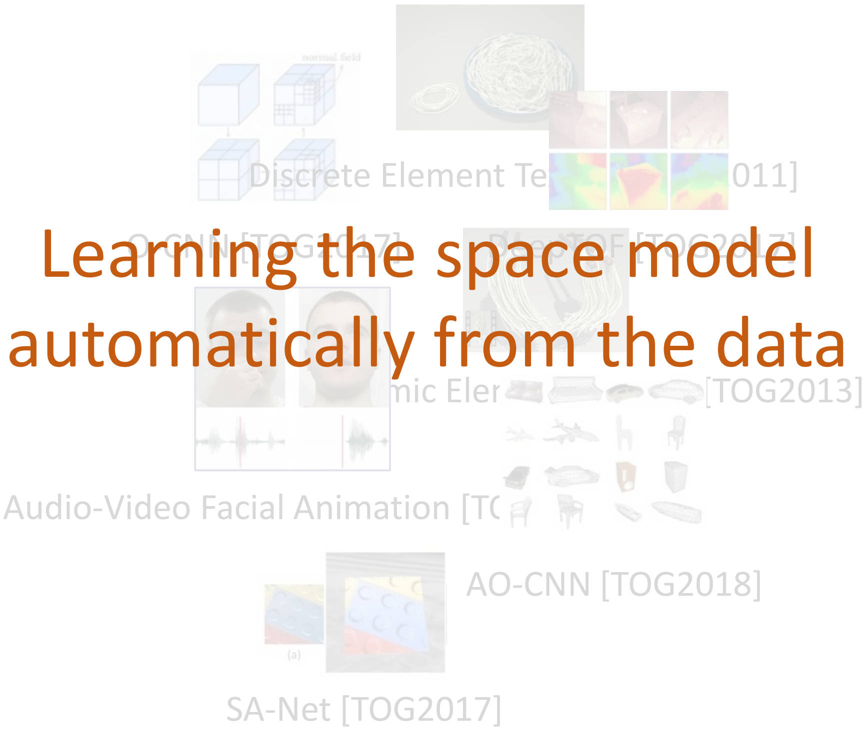
Results: Performance Driven Animation

How to Learn the Model of the Target Space?

Sparse Data

Leveraging the priors of the target space for designing compact space model!

Dense Data



Learning the space model automatically from the data

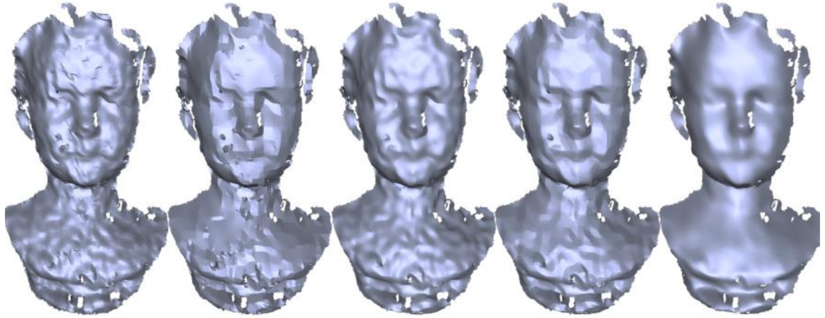
Challenges

- How can I fully utilize the data for model building?

Our Efforts

- How can I fully utilize the data for model building?
- Exploit the representations that can maximize the data coherence

Discrete Element Textures [SIGGRAPH 2011]



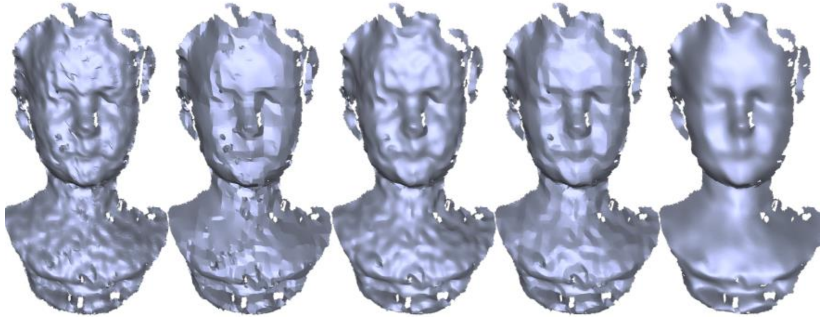
Mesh Denoising with Cascaded Regression [SIgASIA 2016]

Microsoft Research 微软亚洲研究院

Our Efforts

- How can I fully utilize the data for model building?
- Exploit the representations that can maximize the data coherence

Discrete Element Textures [SIGGRAPH 2011]



Mesh Denoising with Cascaded Regression [SIgASIA 2016]

Our Goal

- Automatically generate 3D aggregations from exemplars
  - Different shapes and distributions…
  - From physically plausible to artistic style
  - Easily to edit and manipulate

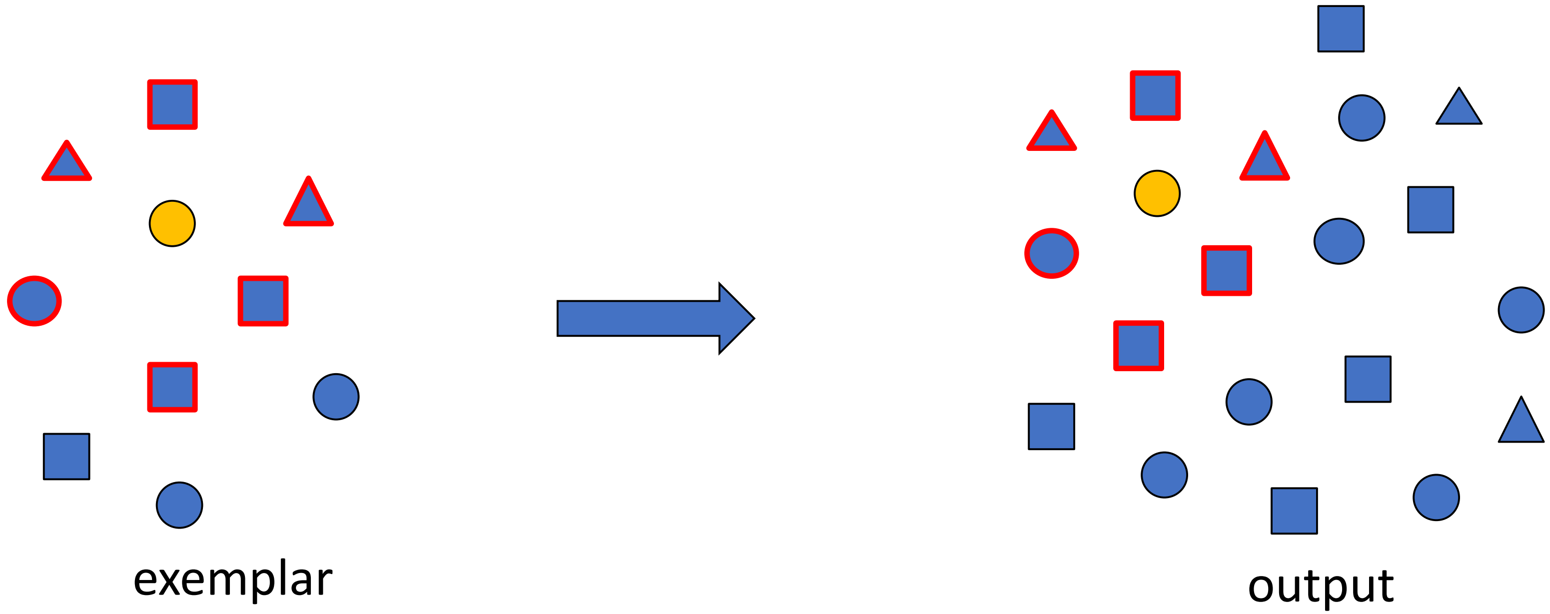




Chong Yang Ma, Li-Yi Wei, Xin Tong, Discrete Element Textures**,** ACM SIGGRAPH 2011

Our Key Observation

- The element distribution follows the Markov random field
  - Each element position is determined by its neighborhood only
- We can learn the local distribution from exemplar directly
  - Copy & paste

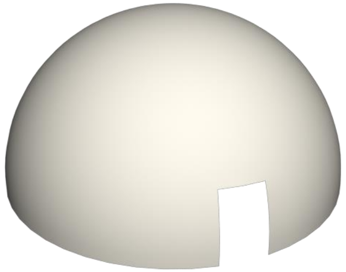


Our Solution

- Extend 2D texture synthesis to discrete elements
  - Non-parametric learning
  - User provides the overall shape and exemplar
  - Algorithm automatically synthesizes the results from exemplar

3D exemplar

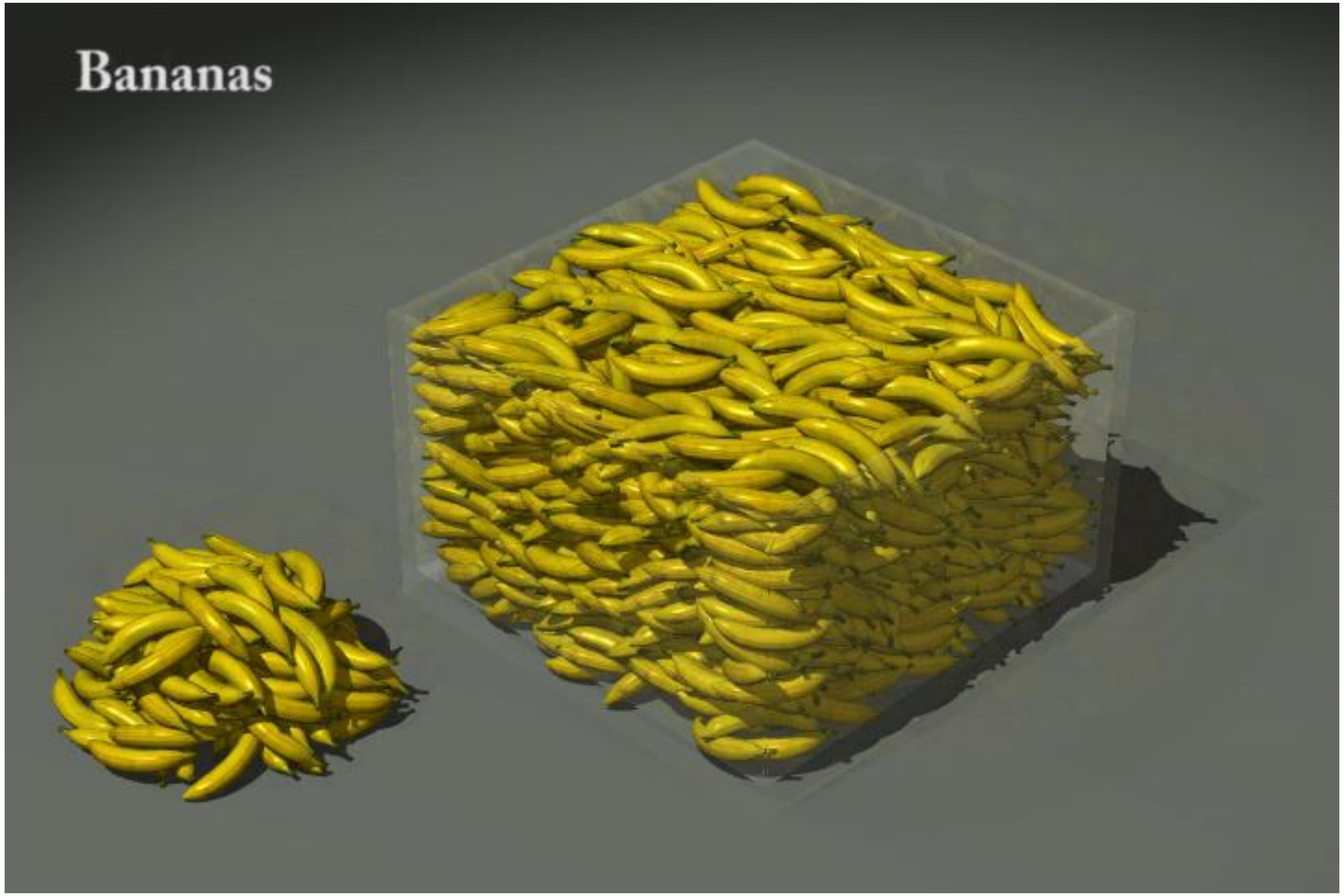
Synthesis



domain

output

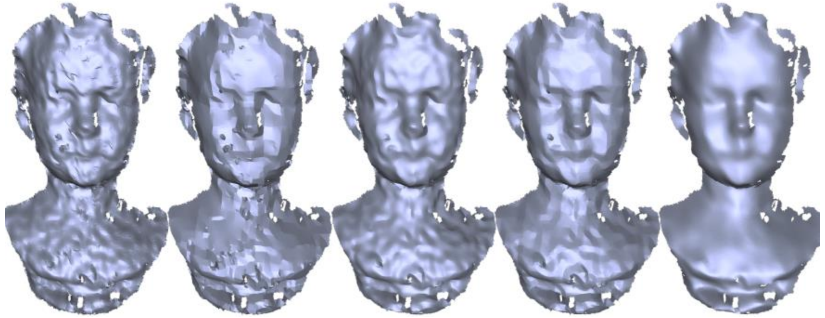
Results



Our Efforts

- How can I fully utilize the data for model building?
- Exploit the representations that can maximize the data coherence

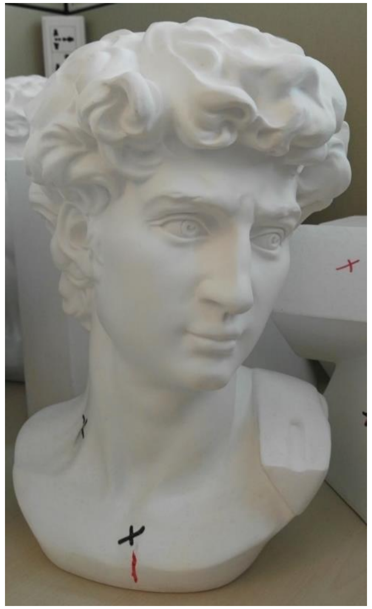
Discrete Element Textures [SIGGRAPH 2011]



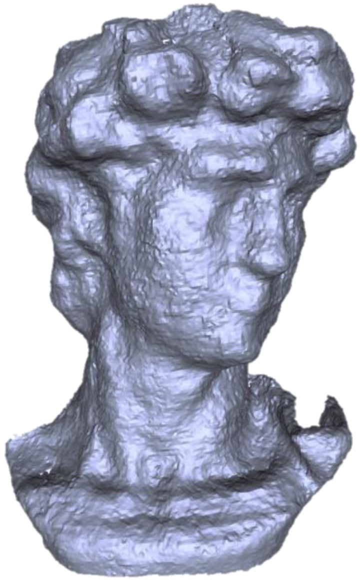
Mesh Denoising with Cascaded Regression
[SIgASIA 2016]

Our Goal

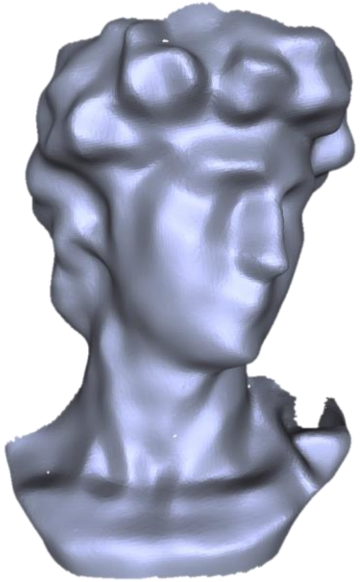
- Removing the noise from scanned 3D mesh
  - Automatic and fast enough



Real object



3D scanning



Denoising result

Pengshuai Wang, Yang Liu, Xin Tong, Mesh Denoising via Cascaded Normal Regressions, ACM SIGGRAPH ASIA 2016

Key Challenges

- Ill-condition problem with unknown ground truth mesh and noise

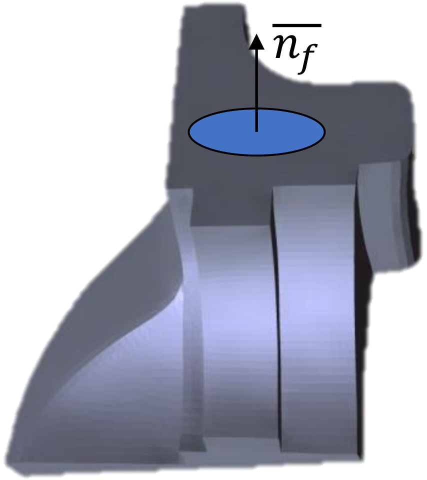
$$M = \overline{M} + \varepsilon$$

- Underline mesh have multi-scale geometry features

- Noise cannot be simple modeled

Key Observations

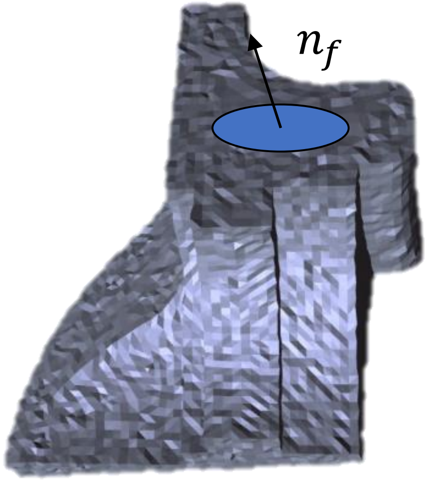
- Normal of a facet can be derived from surrounding facet normal

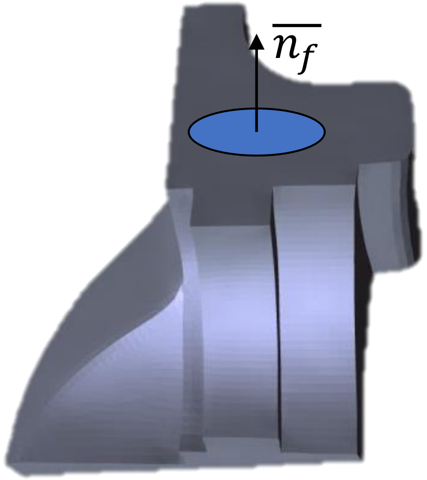


$$\overline{n_f} = G(\, S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \ldots))$$

Key Observations

- Normal of a facet can be derived from surrounding facet normal





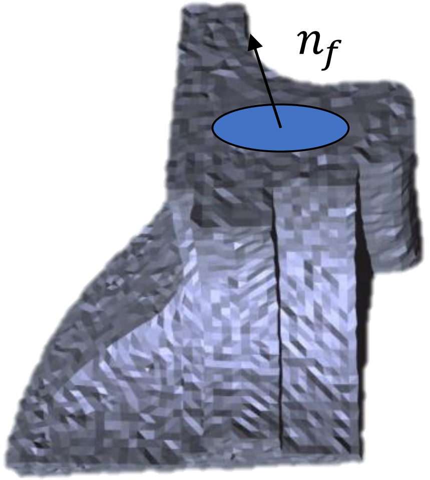
$S(n_{f_1}, n_{f_2}, n_{f_3} \dots) \sim S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \dots)$

$\overline{n_f} = G(\, S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \dots))$

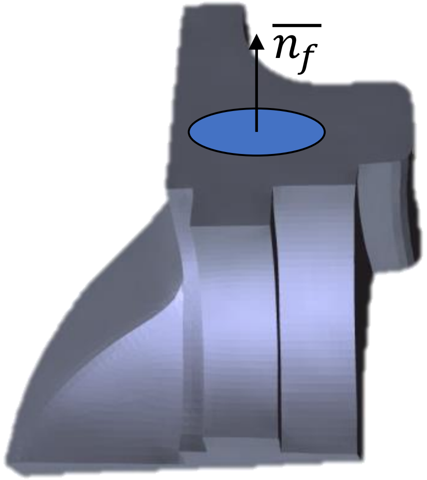
Microsoft Research 微软亚洲研究院

Key Observations

- Normal of a facet can be derived from surrounding facet normal



$$\overline{n_f} = G'(\, S(n_{f_1}, n_{f_2}, n_{f_3} \ldots))$$

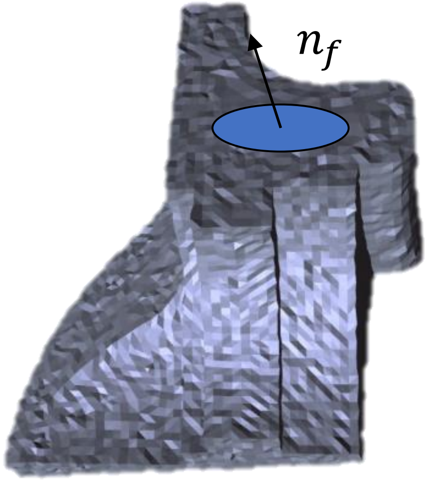


$$S(n_{f_1}, n_{f_2}, n_{f_3} \ldots) \sim S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \ldots)$$

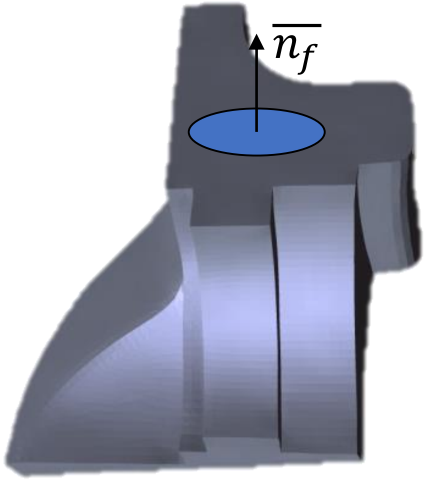
$$\overline{n_f} = G(\, S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \ldots))$$

Key Observations

- Normal of a facet can be derived from surrounding facet normal
- We can learn the function G' from a set of mesh pairs



$$\overline{n_f} = G'(\,S(n_{f_1}, n_{f_2}, n_{f_3} \dots))$$

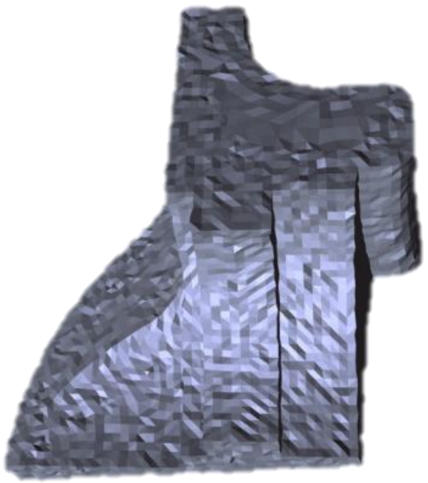


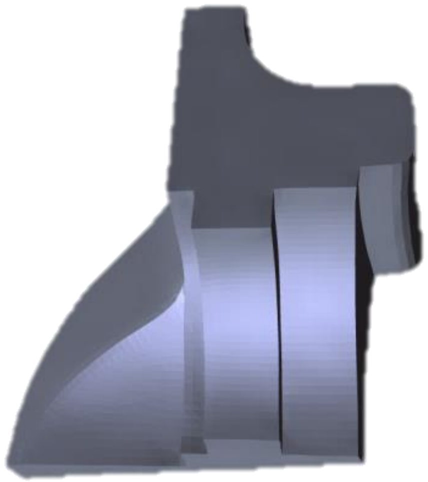
$$S(n_{f_1}, n_{f_2}, n_{f_3} \dots) \sim S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \dots)$$

$$\overline{n_f} = G(\,S(\overline{n_{f_1}}, \overline{n_{f_2}}, \overline{n_{f_3}} \dots))$$

Our Solution

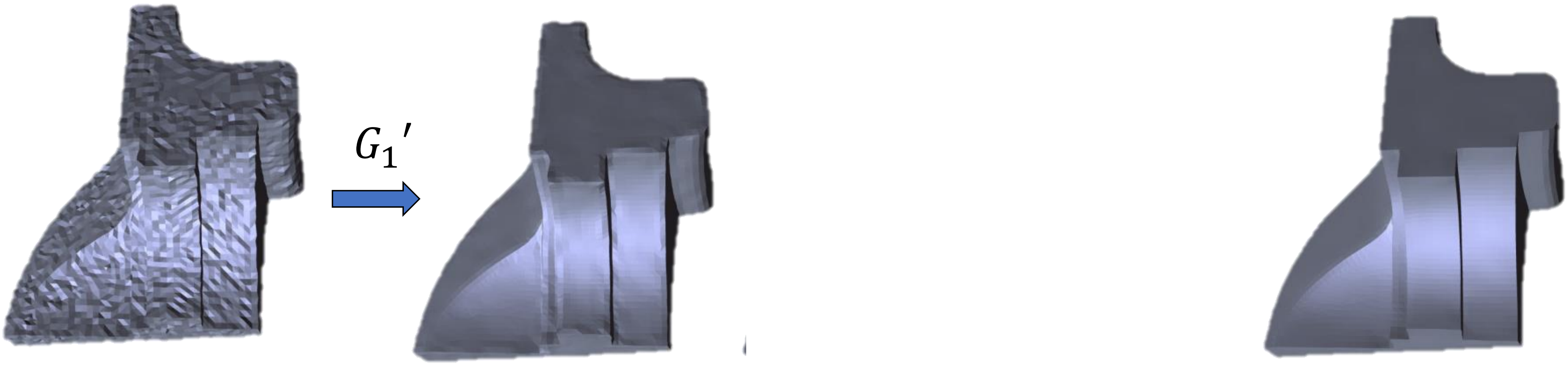
- Define a set of bi-lateral normal filter results as features S
  - Filtered facet normal descriptor (FND)
- Learn the function G' with cascaded regression functions
  - RBF neural networks as regression function in each step





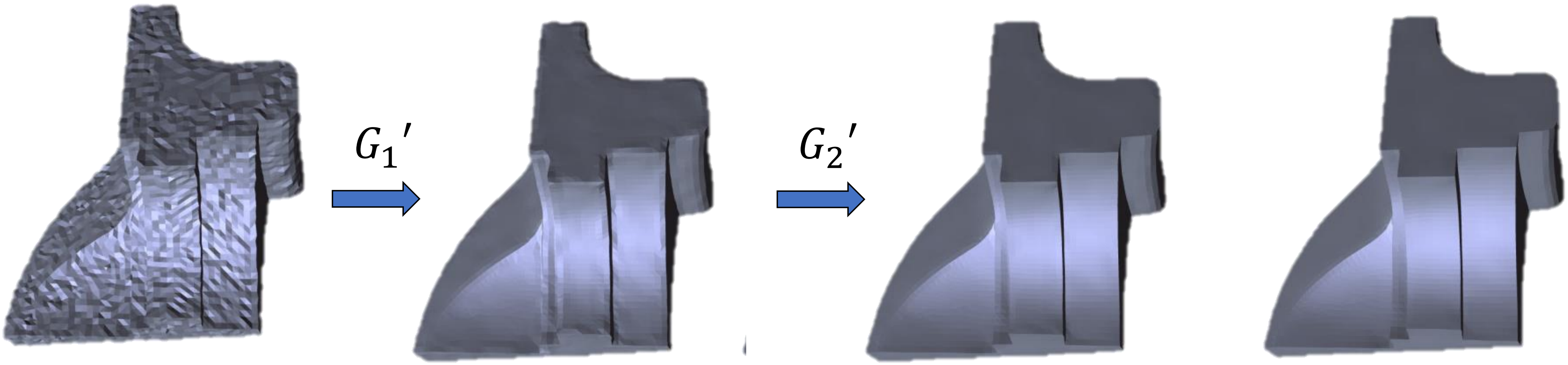
Our Solution

- Define a set of bi-lateral normal filter results as features S
    - Filtered facet normal descriptor (FND)
- Learn the function G' with cascaded regression functions
    - RBF neural networks as regression function in each step



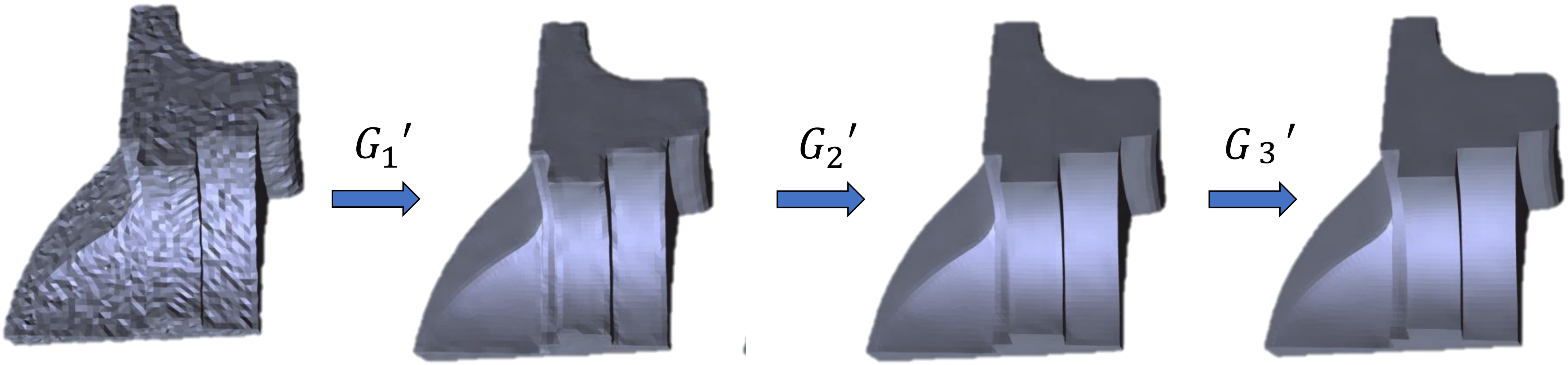
Our Solution

- Define a set of bi-lateral normal filter results as features S
  - Filtered facet normal descriptor (FND)
- Learn the function G' with cascaded regression functions
  - RBF neural networks as regression function in each step



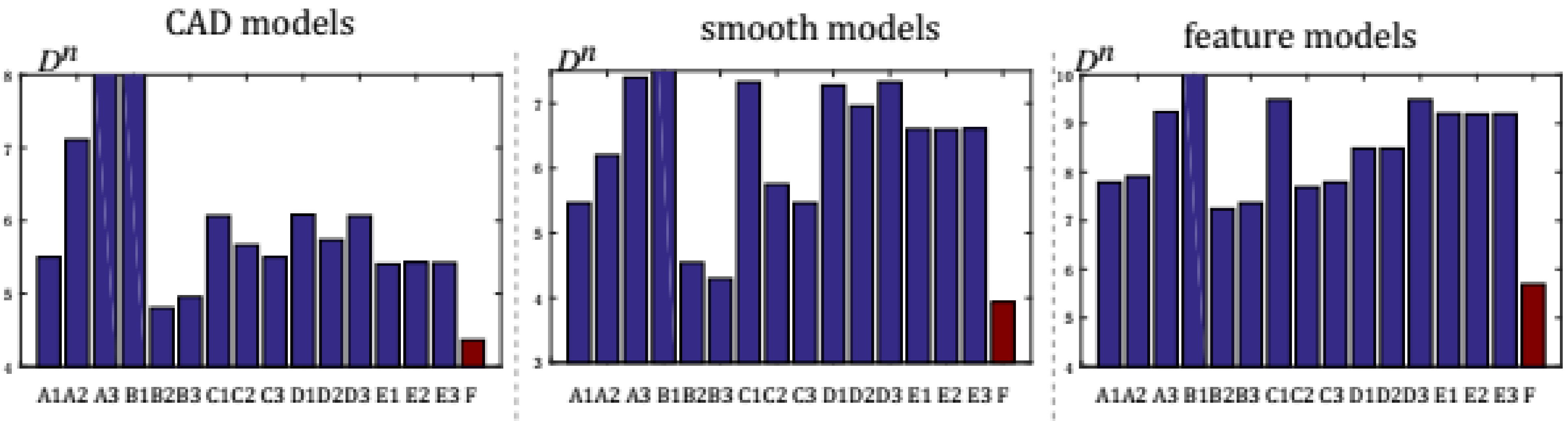
Our Solution

- Define a set of bi-lateral normal filter results as features S
  - Filtered facet normal descriptor (FND)
- Learn the function G' with cascaded regression functions
  - RBF neural networks as regression function in each step



Microsoft Research 微软亚洲研究院

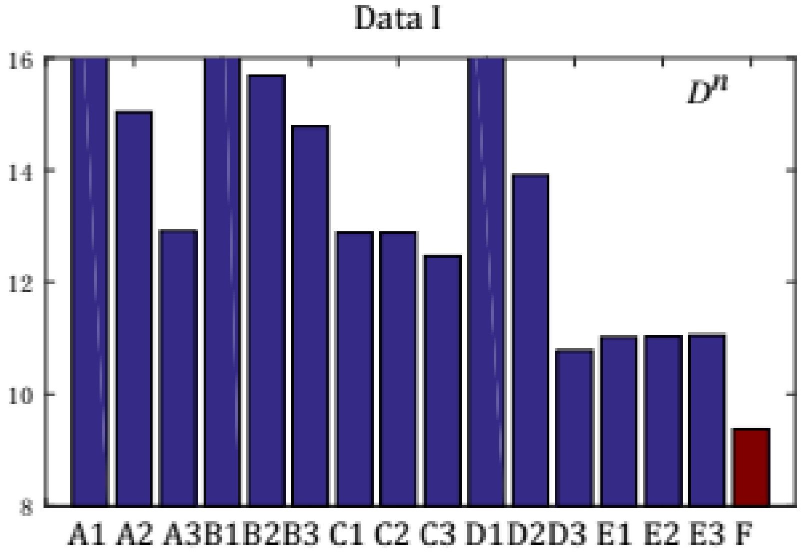
Results: Synthetic Data

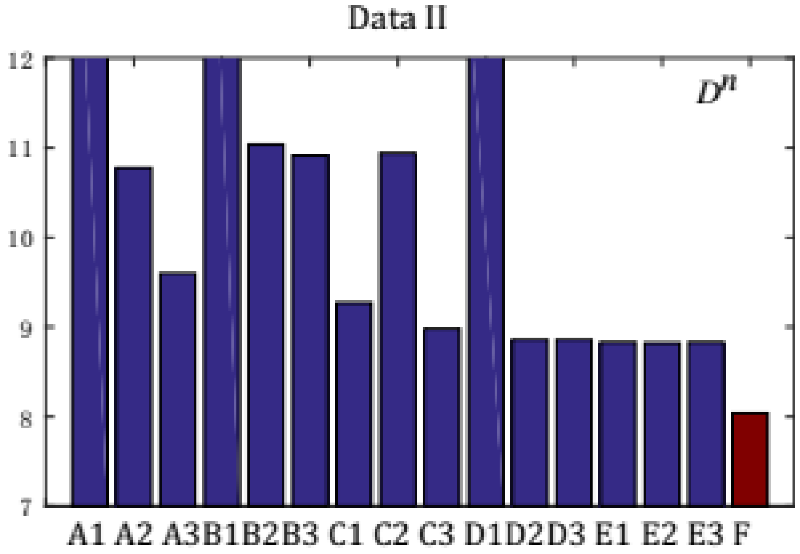


Results: Synthetic Data

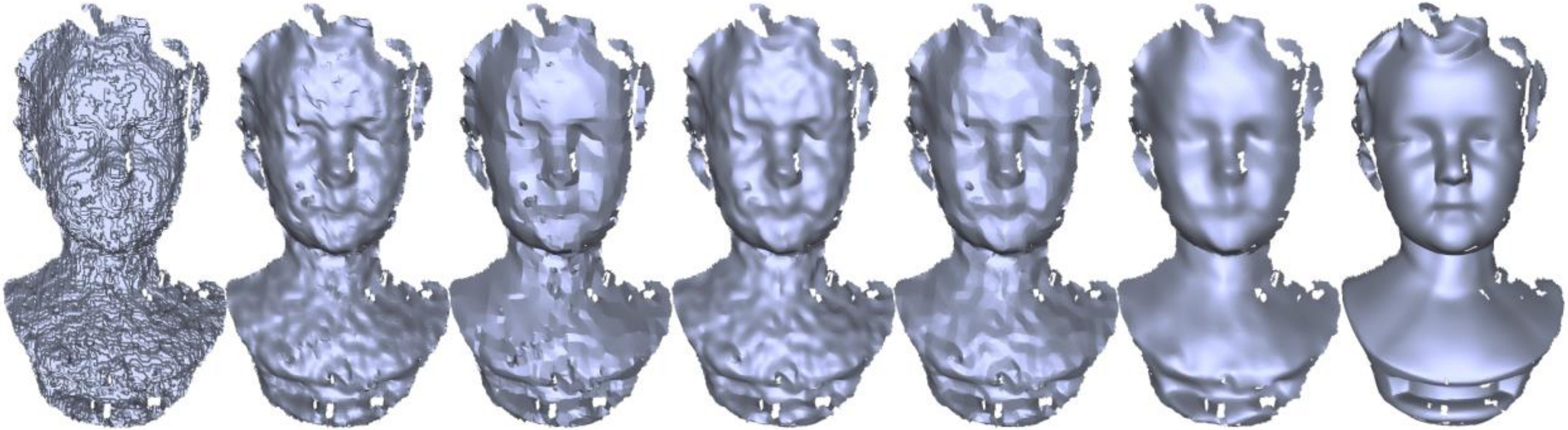
(a) Noisy input (b) Bilateral normal (c) L_0 smoothing (d) Guided normal (e) Bayesian (f) Our method (g) Ground-truth

Results: Real Data



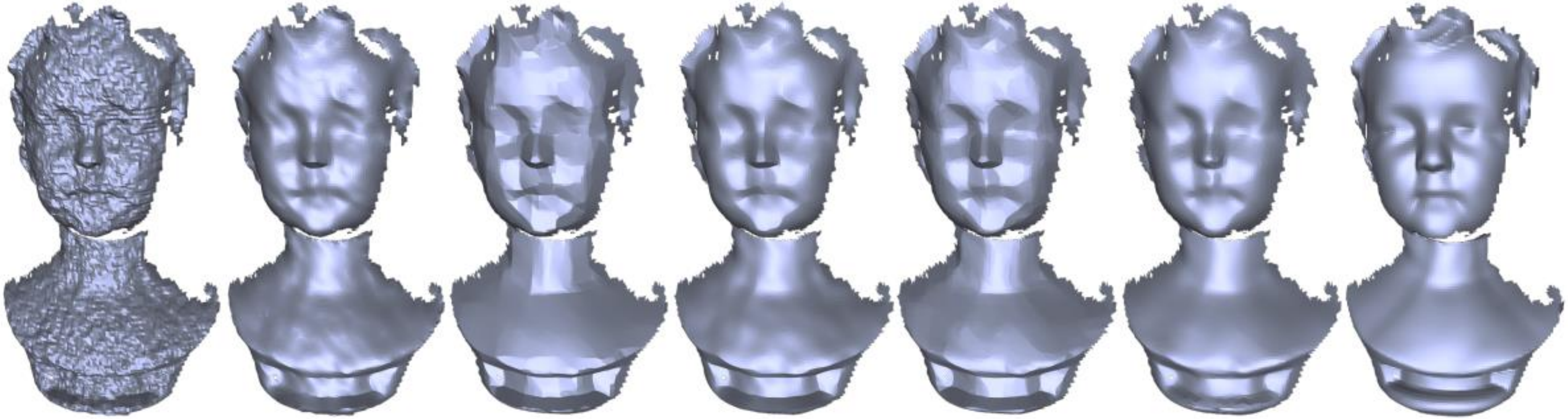


Results: Real Data with Kinect V1



(a) Noisy input (b) Bilateral normal (c) L_0 smoothing (d) Guided normal (e) Bayesian (f) Our method (g) Ground-truth

Results: Real Data with Kinect V2



(a) Noisy input (b) Bilateral normal (c) L_0 smoothing (d) Guided normal (e) Bayesian (f) Our method (g) Ground-truth

Microsoft Research 微软亚洲研究院

Results: Performance

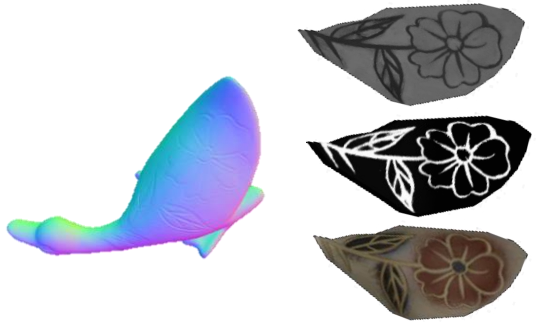
N_f	10k	25k	54k	99k	171k	566k
Bilateral normal	1.2s	2.7s	6.3s	14.2s	23.7s	71.4s
L_0 smoothing	4.7s	37.1s	286.2s	622.4s	885.2s	3155.7s
Guided normal	2.6s	7.2s	19.2s	44.9s	99.7s	558.9s
Bayesian	6.1s	16.5s	39.1s	76.6s	126.2s	394.6s
Our method	**0.8**s	**1.8**s	**2.9**s	**5.7**s	**11.3**s	**28.3**s

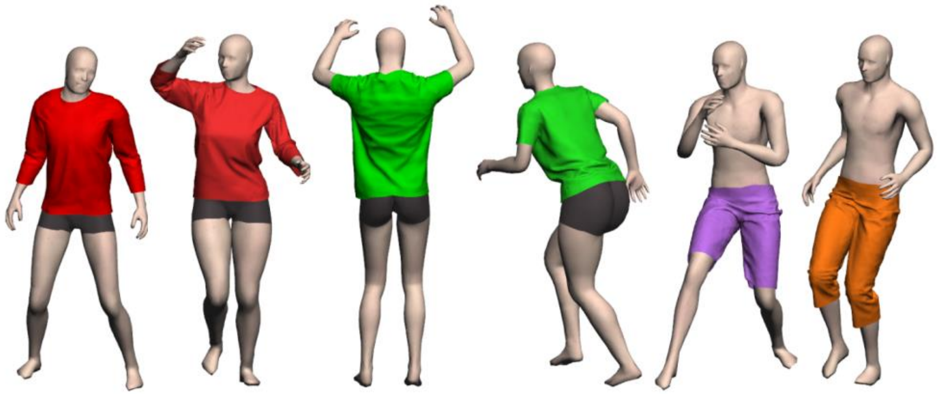
Outline

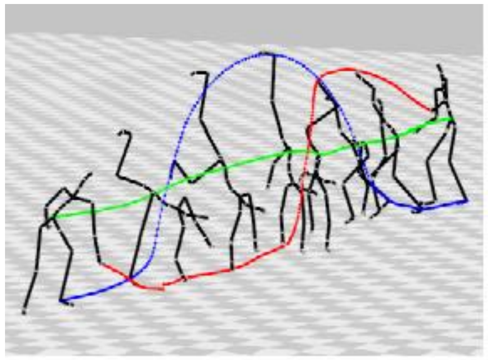
- An overview of data driven graphics
- Key challenges in data driven graphics and our exploration
- Future directions

Fundamental Challenges

- High dimensionality of the graphics functions and data
  - Geometry, appearance, dynamics, and their interactions (light transport)

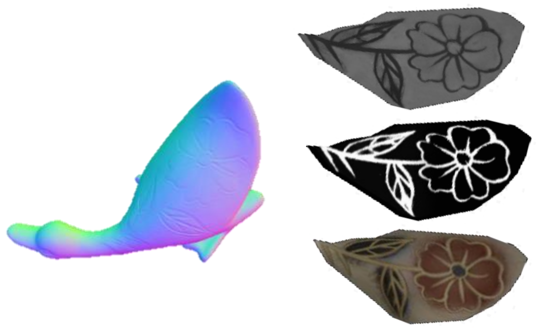


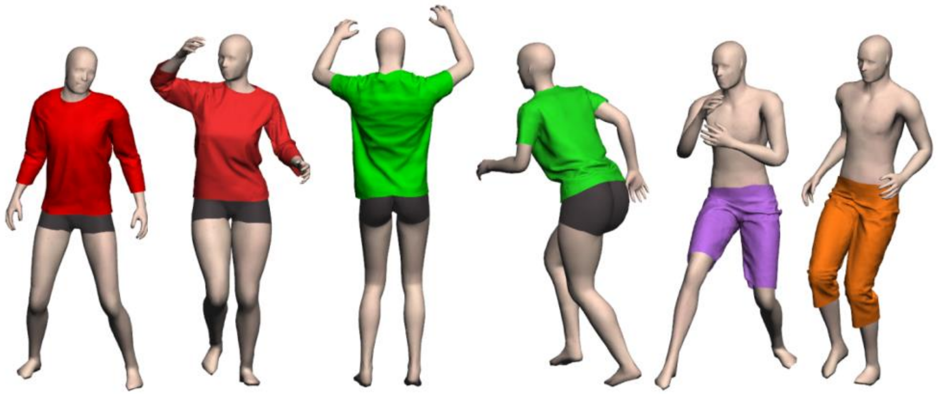


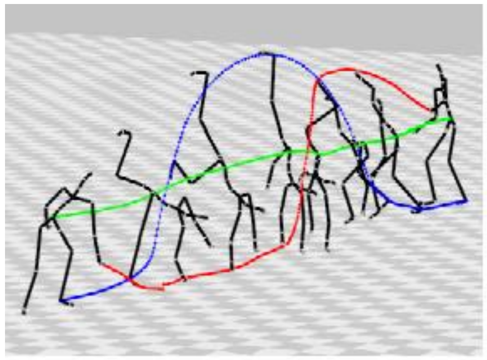


Key Challenges

- High dimensionality of the graphics functions and data
  - Geometry, appearance, dynamics, and their interactions (light transport)
  - Data is difficult to be acquired and measured (small labeled dataset)
  - Dimensionality gap between the data and observation (image/video)
  - Variant representations and measurements

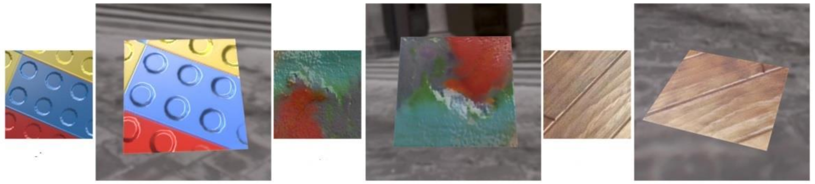






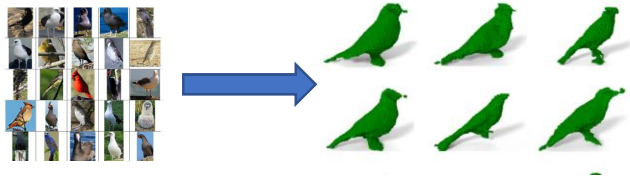
Our Efforts

Small labeled dataset



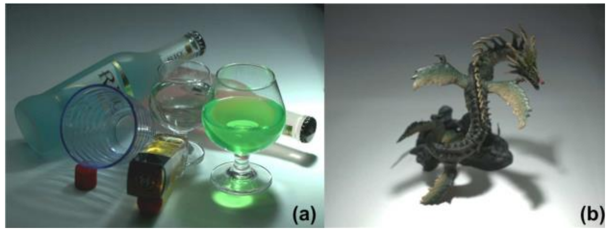
Self-augmented CNN training for SVBRDF modeling [SIGGRAPH 2017]

Dimensionality gap



Multi-projection GAN [CVPR 2019]

Variant representations

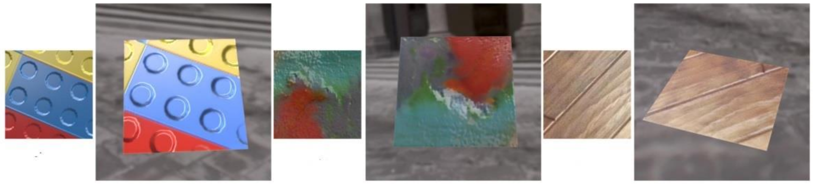


Kernel Nystorm for Relighting [SIGGRAPH 2009]

Image based Relighting [SIGGRAPH 2015]

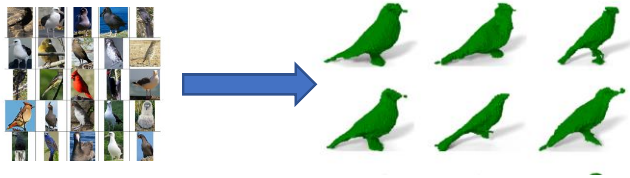
Our Efforts

Small labeled dataset



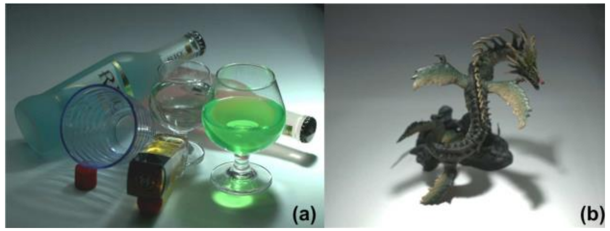
Self-augmented CNN training for SVBRDF modeling [SIGGRAPH 2017]

Dimensionality gap



Multi-projection GAN [CVPR 2019]

Variant representations

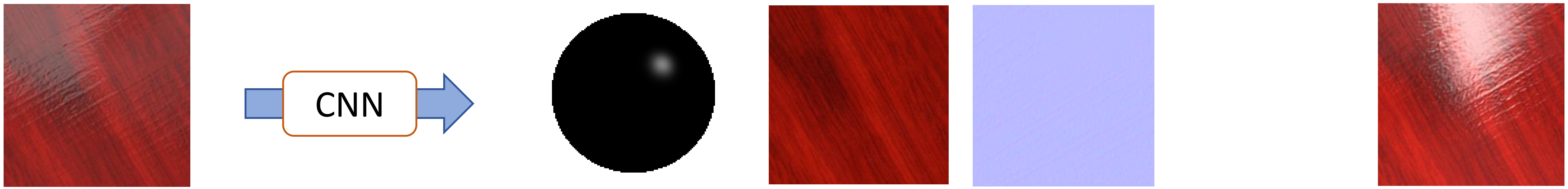


Kernel Nystorm for Relighting [SIGGRAPH 2009]

Image based Relighting [SIGGRAPH 2015]

Our Goal

- Material modeling from a single image using CNN
  - Replace tedious manual work done by skilled artist
  - Automatic and fast
  - Reasonable quality

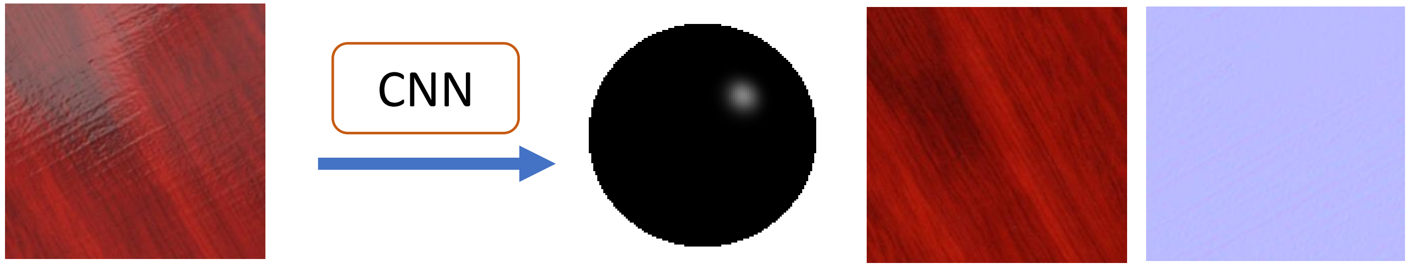


Input image | Specular | Diffuse map | Normal map | New rendering

Surface material map

Xiao Li, Yue Dong, Pieter Peers, Xin Tong, *Modeling Surface Appearance from a Single Photograph using Self-Augmented Convolutional Neural Networks*, ACM Transactions on Graphics(SIGGRAPH), 36(4), 2017.

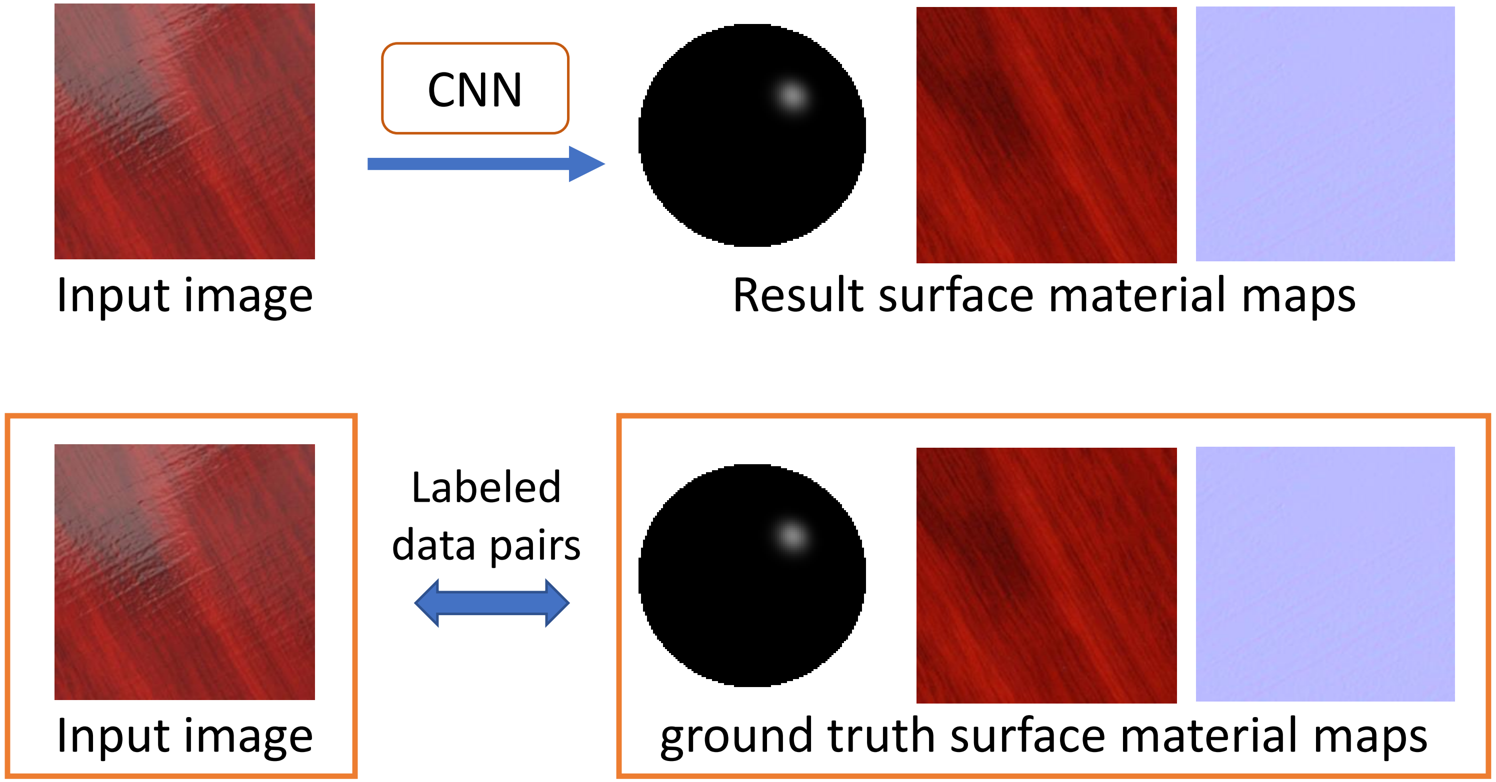
Key Challenge



Input image

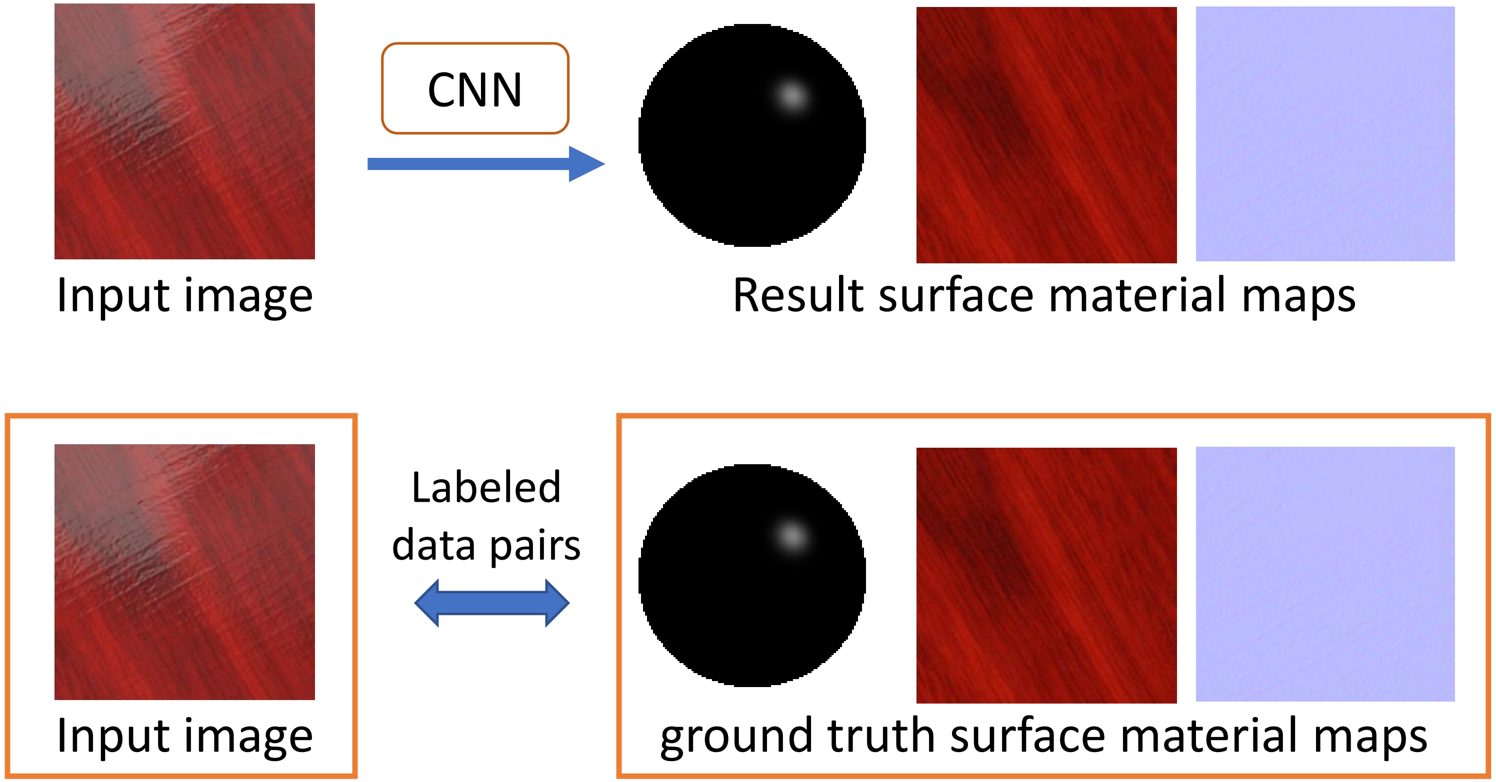
Result surface material maps

Key Challenge



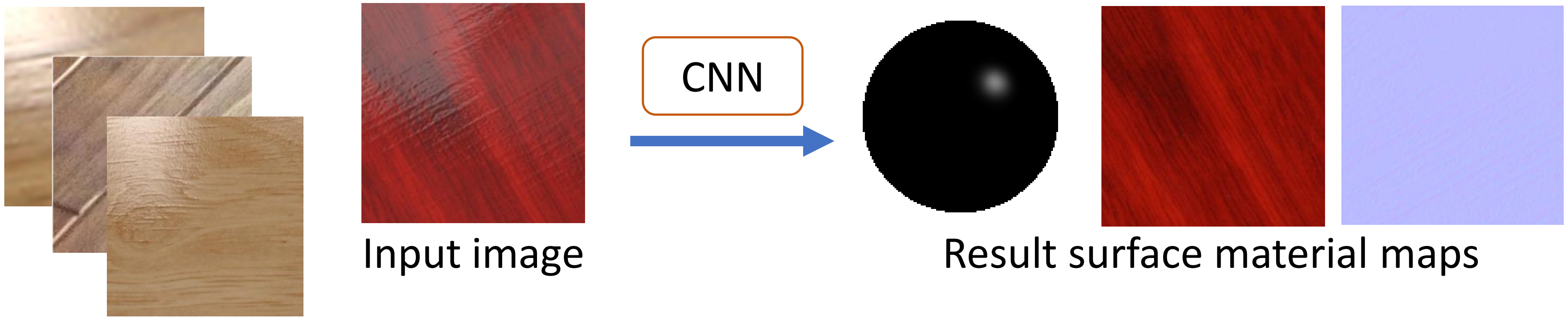
Key Challenge

- We do not have sufficient labeled data for training



Our Key Observations

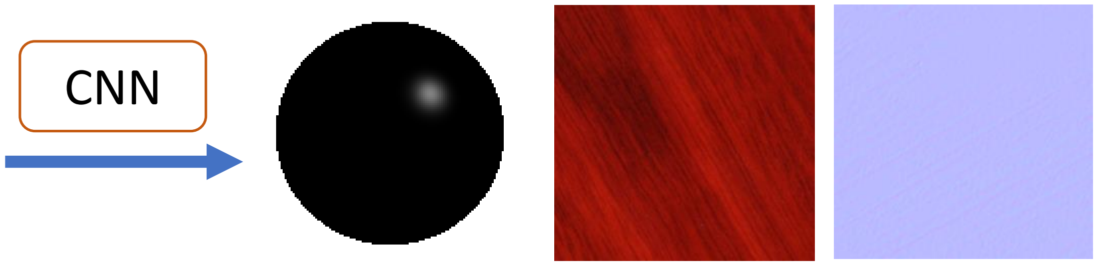
- We do have large amount of unlabeled images



Our Key Observations

- We do have large amount of unlabeled images

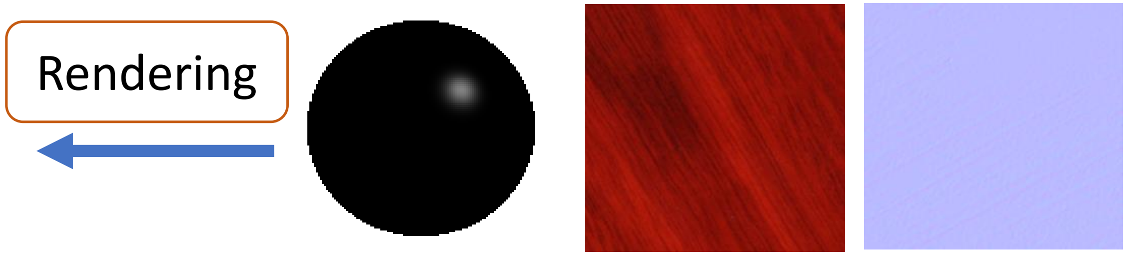
Input image



Result surface material maps

- Inverse mapping of CNN is known: rendering

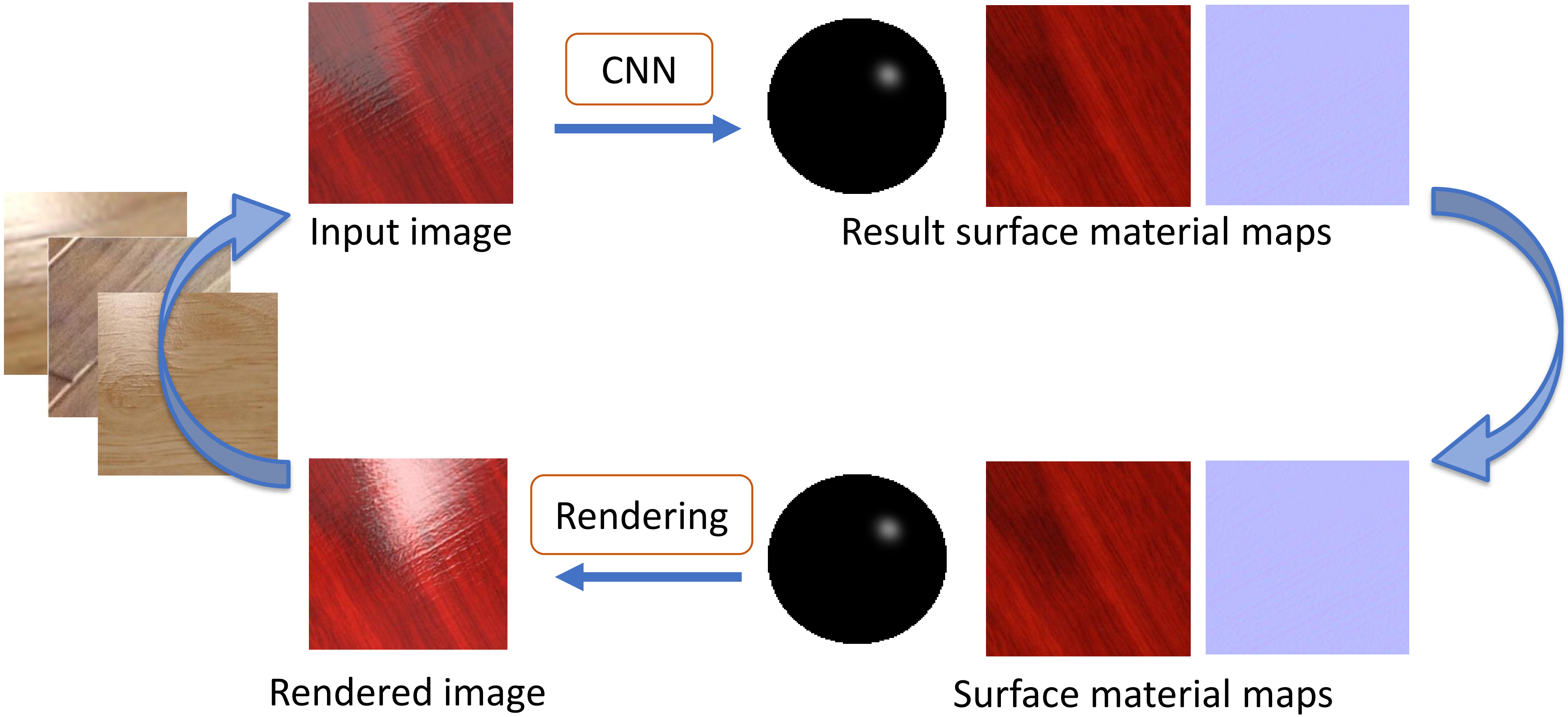
Rendered image



Surface material maps

Our Solution: Self-Augmented CNN Training

- Training CNN with labeled/unlabeled data with the help of rendering



Self-Augmented CNN

Training with labeled data

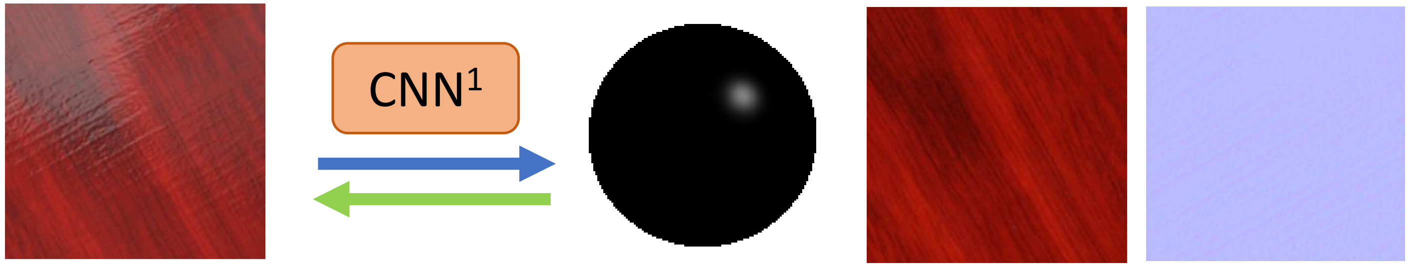
Input image

CNN^0

Ground truth surface material maps

Self-Augmented CNN

Training with labeled data

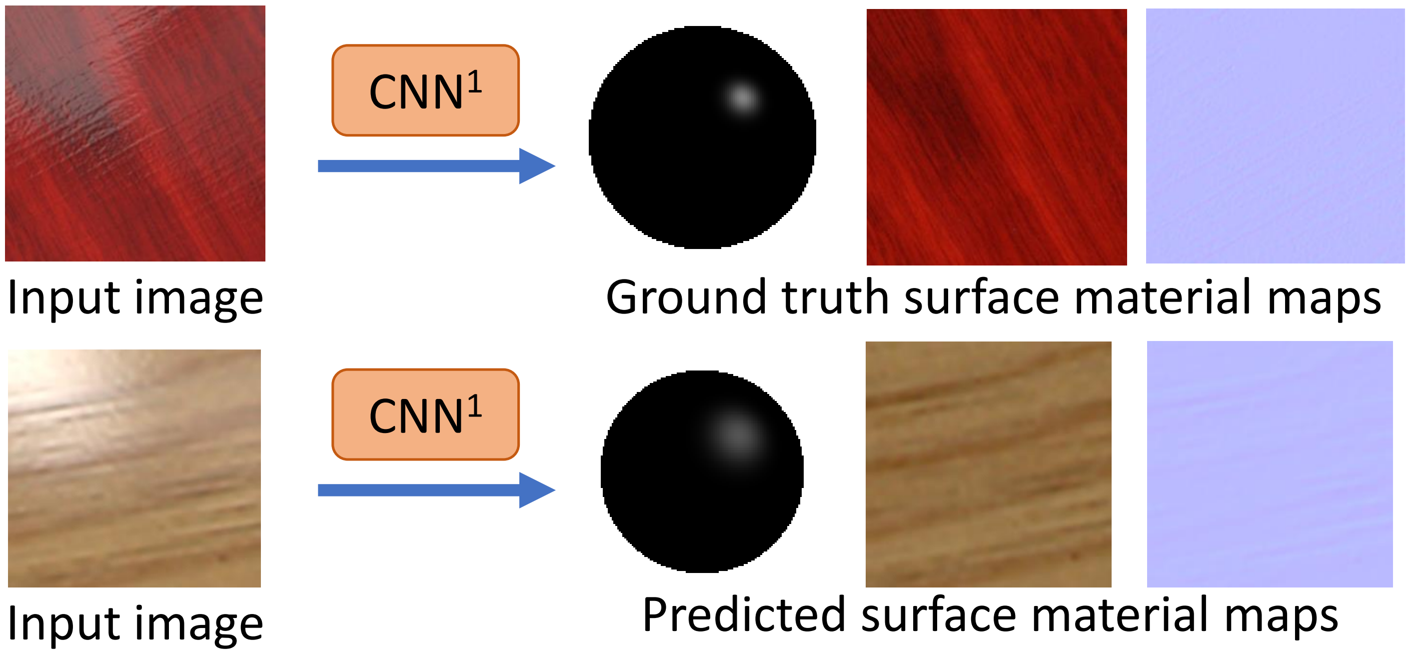


Input image

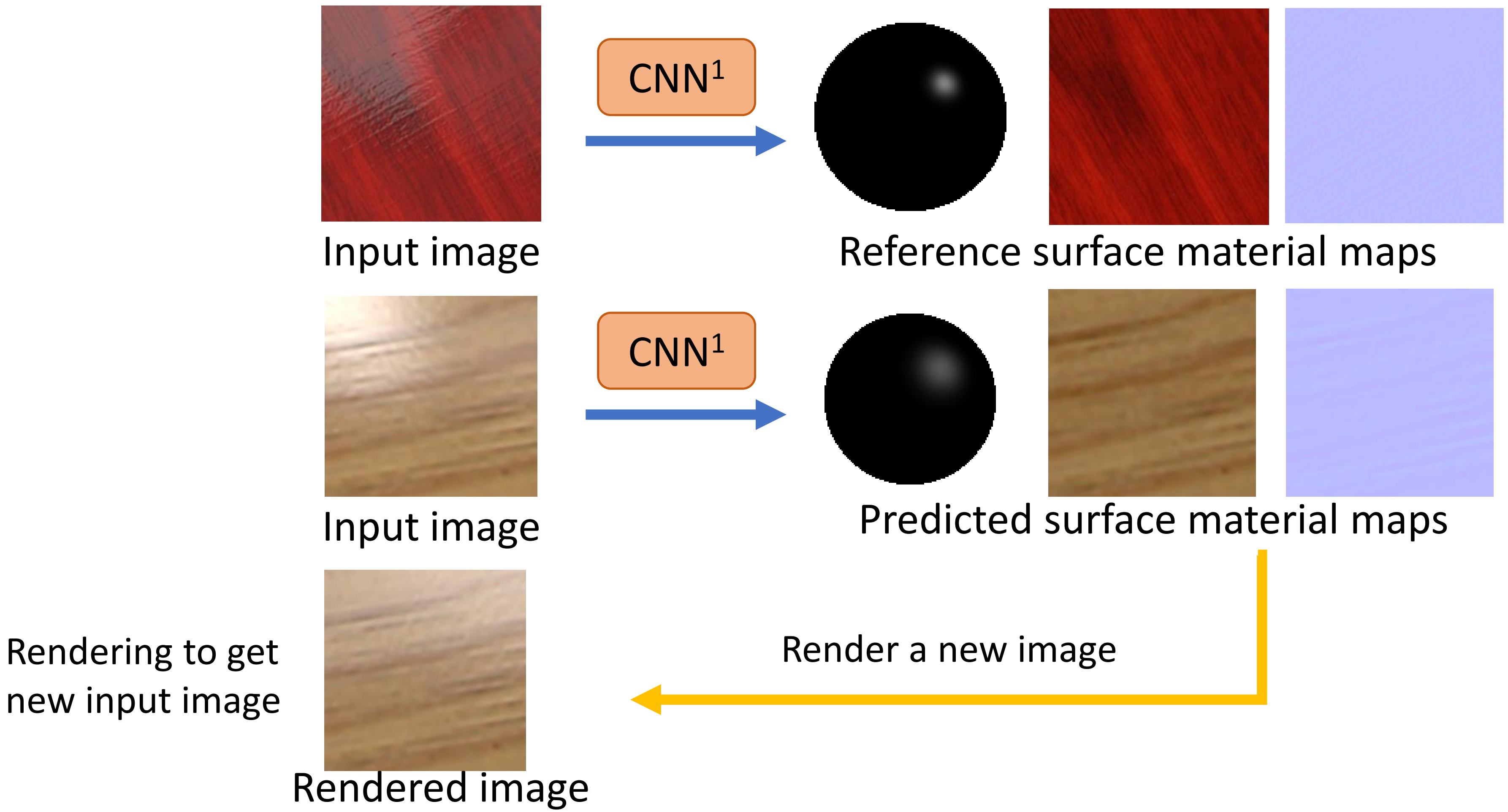
Ground truth surface material maps

Self-Augmented CNN

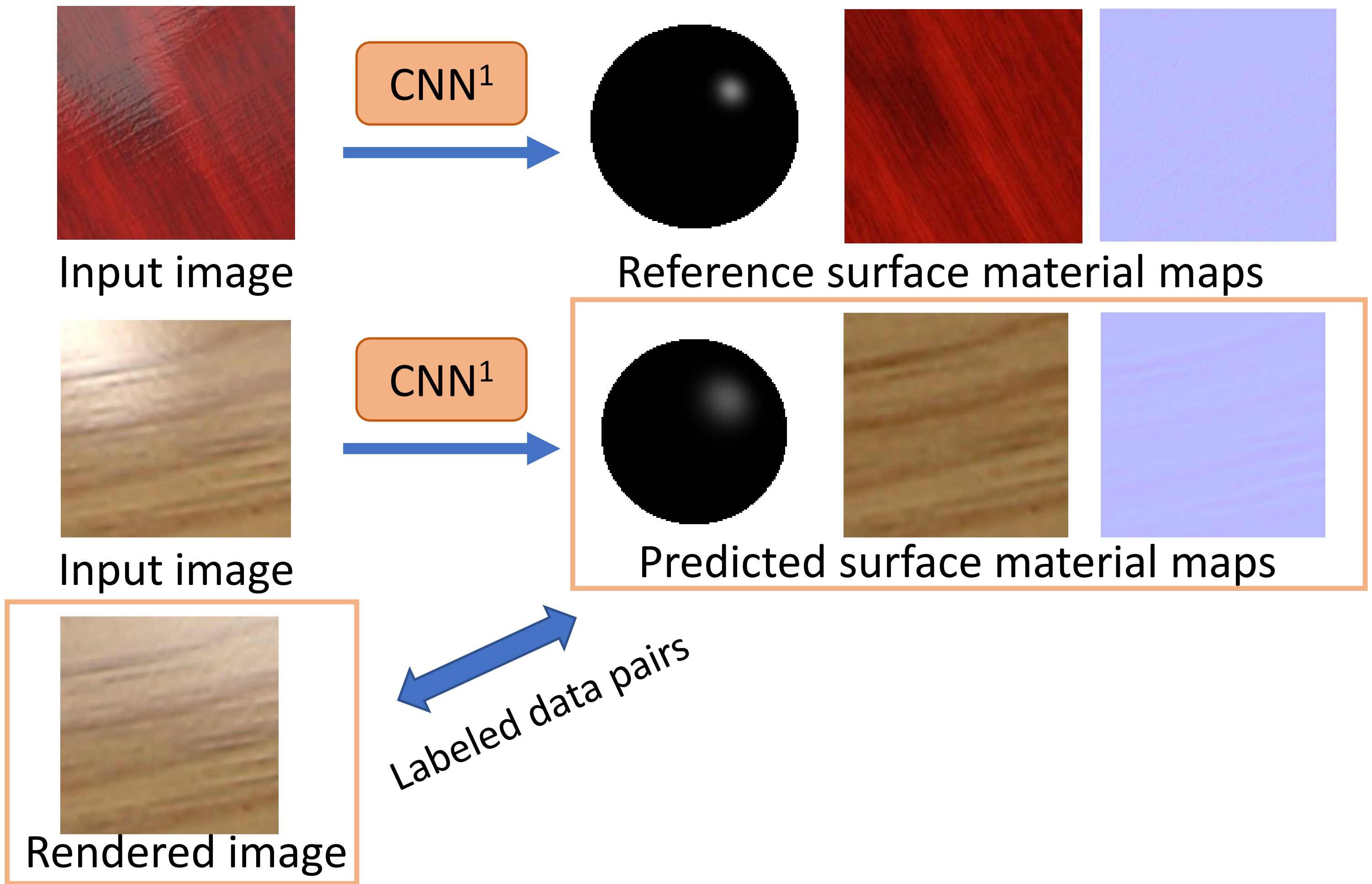
Prediction for unlabeled data



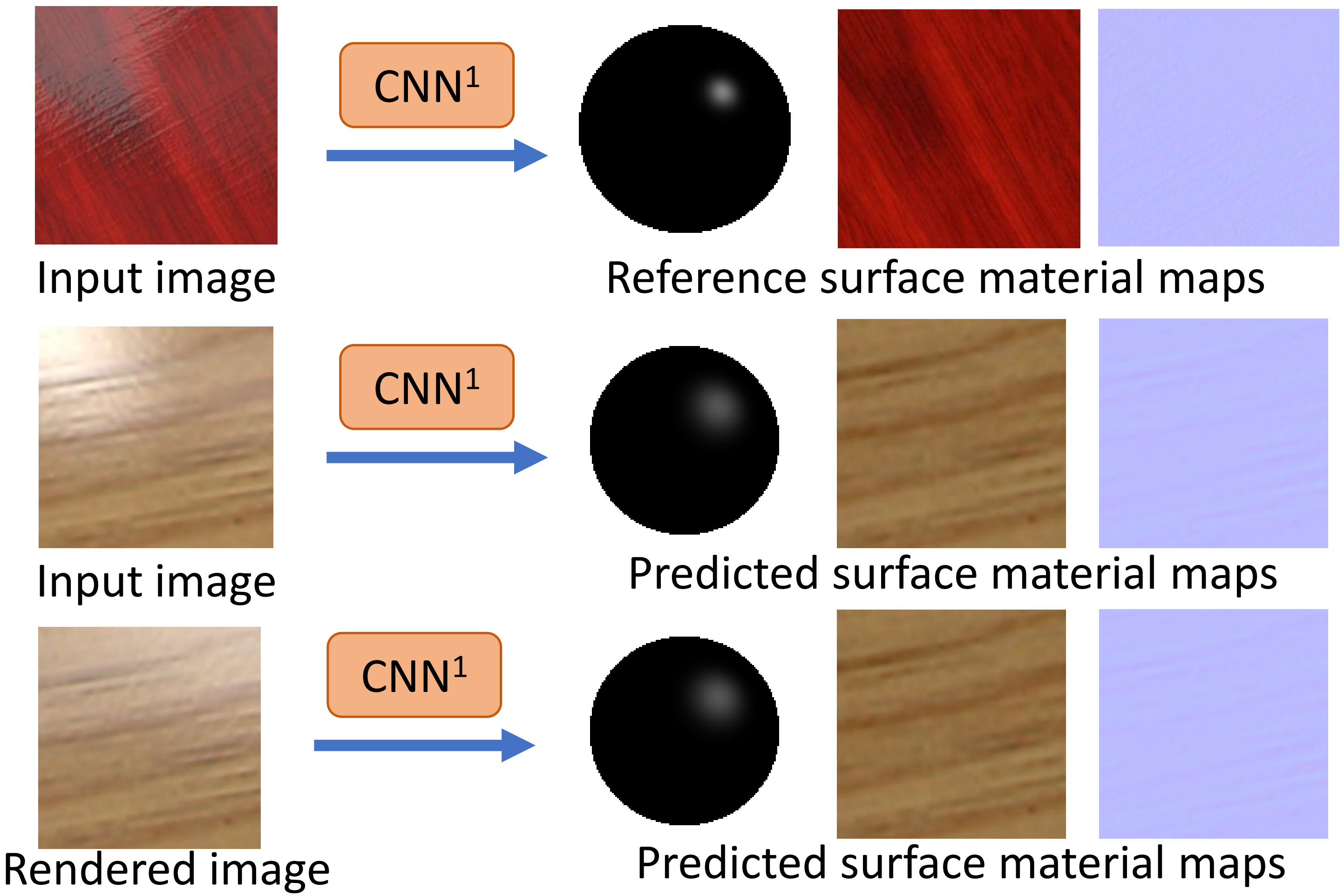
Self-Augmented CNN



Self-Augmented CNN

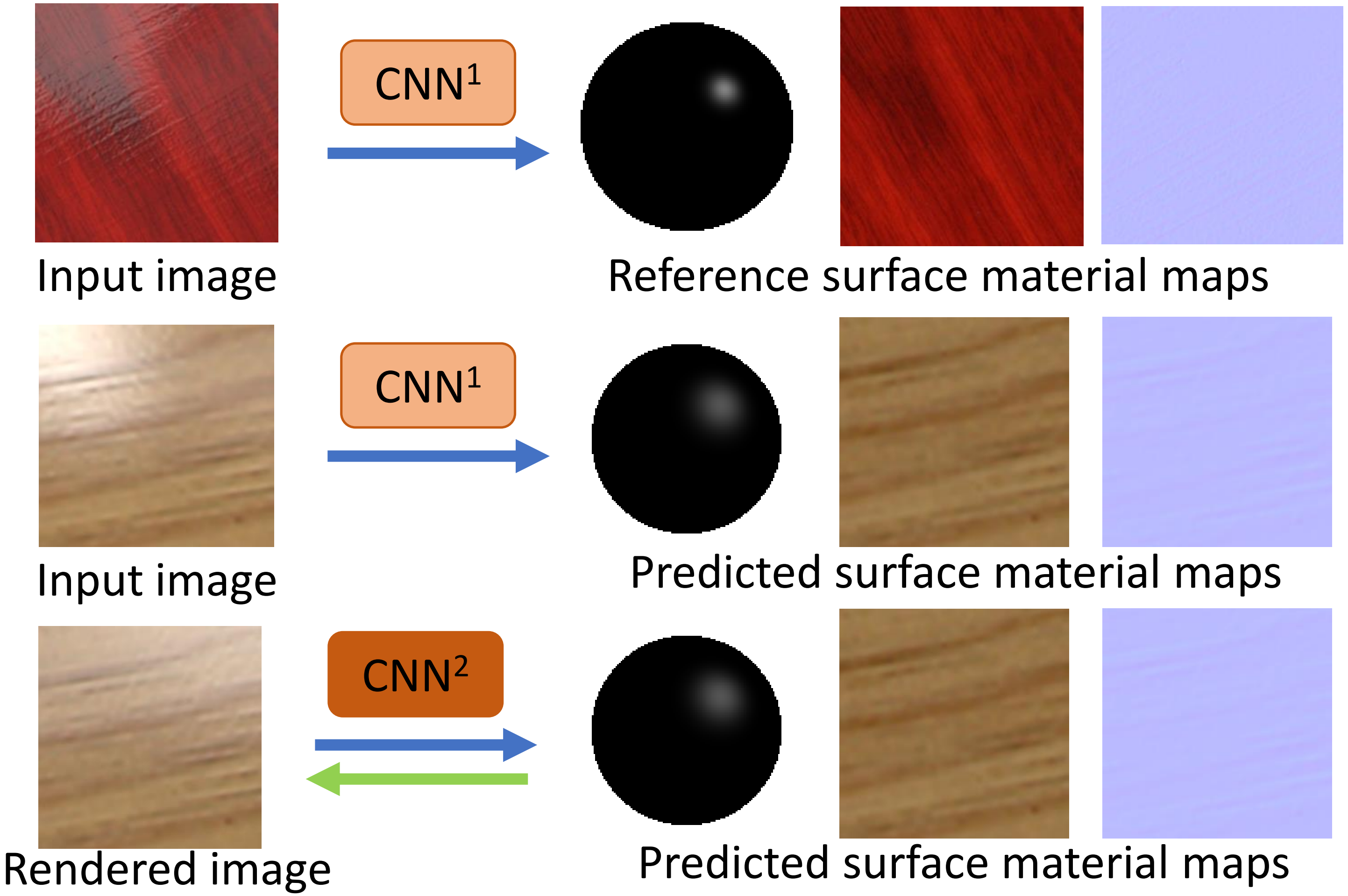


Self-Augmented CNN

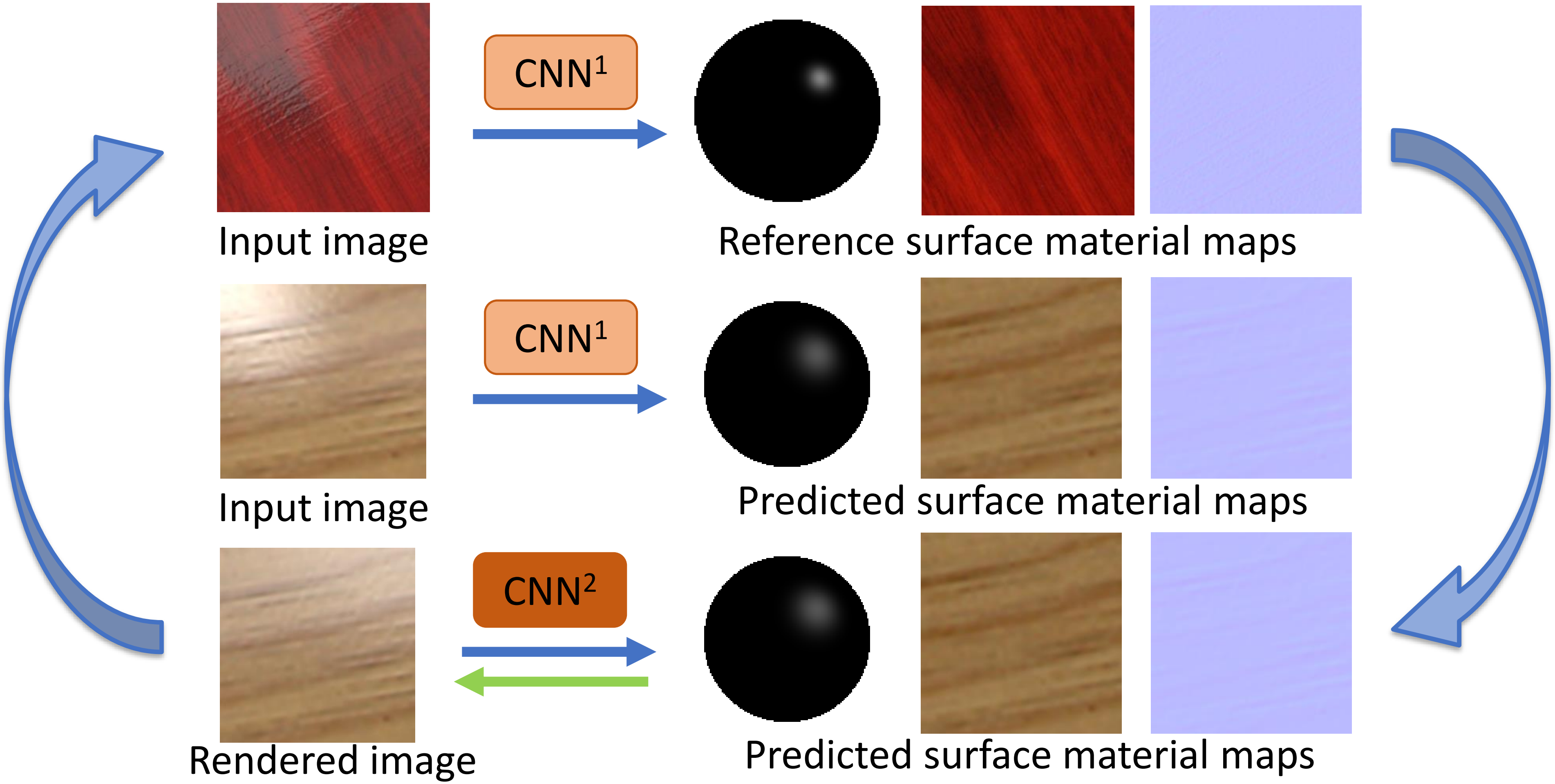


Training with
rendered pairs

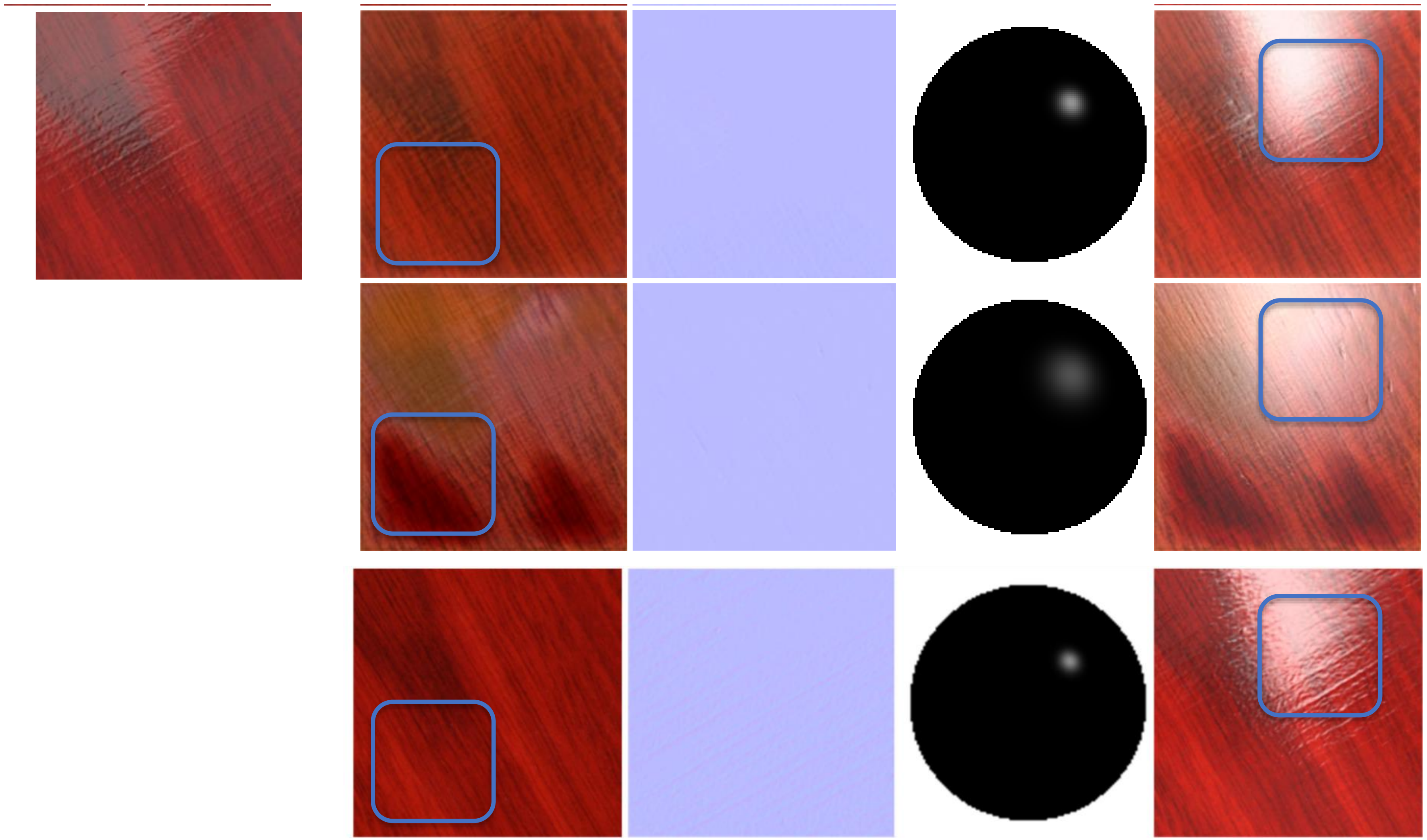
Self-Augmented CNN



Self-Augmented CNN



Comparisons

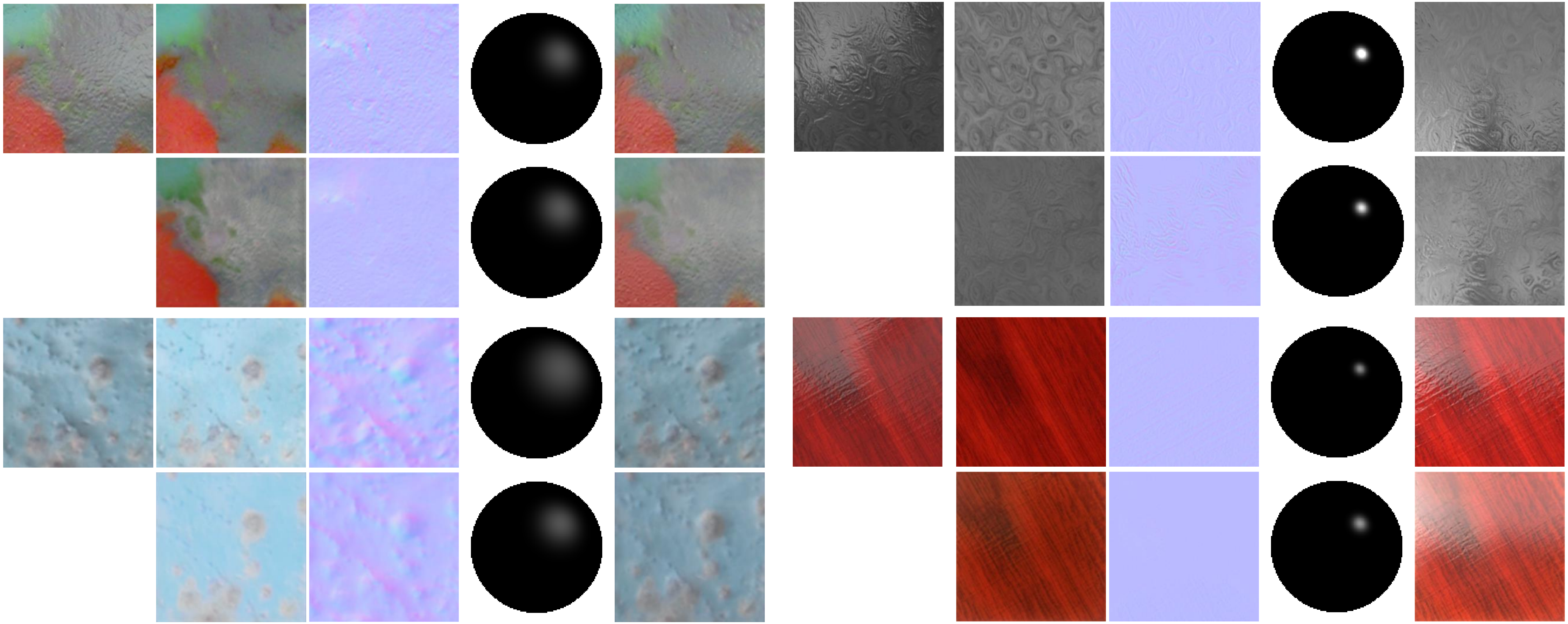


With Self-Augmented Training

Without Self-Augmented Training

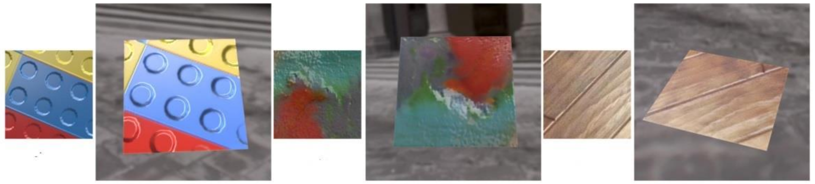
Ground truth

Results



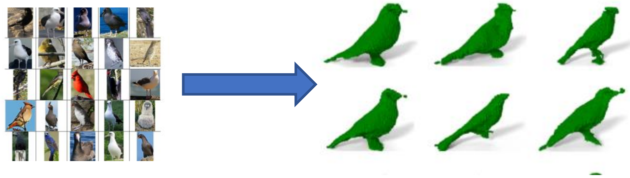
Our Efforts

Small labeled dataset



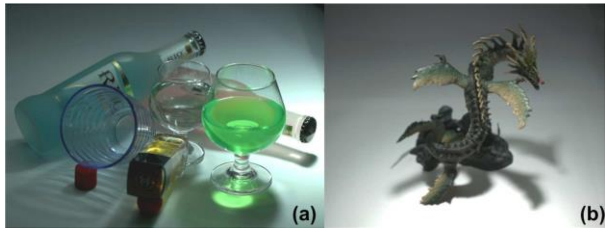
Self-augmented CNN training for SVBRDF modeling [SIGGRAPH 2017]

Dimensionality gap



Multi-projection GAN [CVPR 2019]

Variant representations

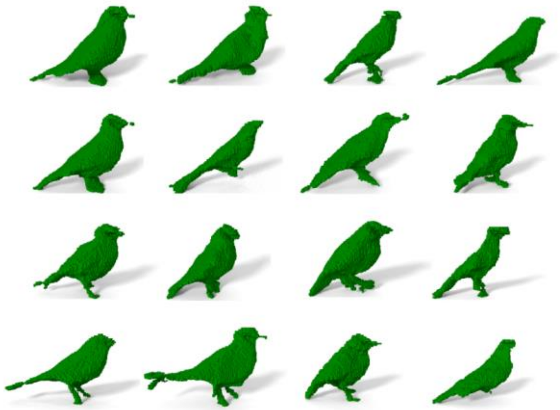


Kernel Nystorm for Relighting [SIGGRAPH 2009]

Image based Relighting [SIGGRAPH 2015]

Our Goal

- Generating 3D shapes (high dimensional) from unannotated 2D image (low dimensional projection) collections
  - Input : 2D silhouettes of the objects captured from different views
  - Output: 3D shapes of the objects in the same class



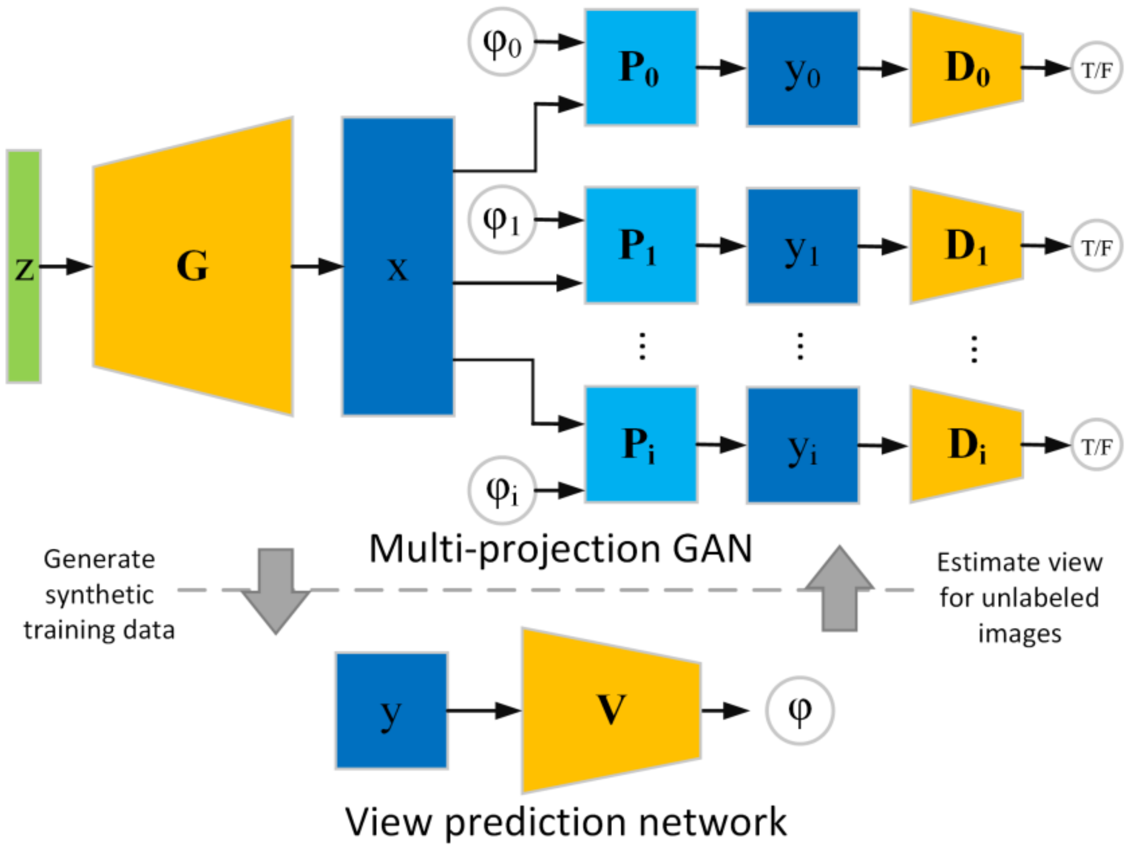
Xiao Li, Yue Dong, Pieter Peers, Xin Tong, *Synthesizing 3D Shapes from Unannotated Image Collections using Multi-projection Generative Adversarial Networks* , Accepted by CVPR 2019.

The Key Challenges

- Gap between 2D image and 3D shapes
- Image has no correspondence
  - We don’t have multiple view images of one object
- View information of each image is unknown

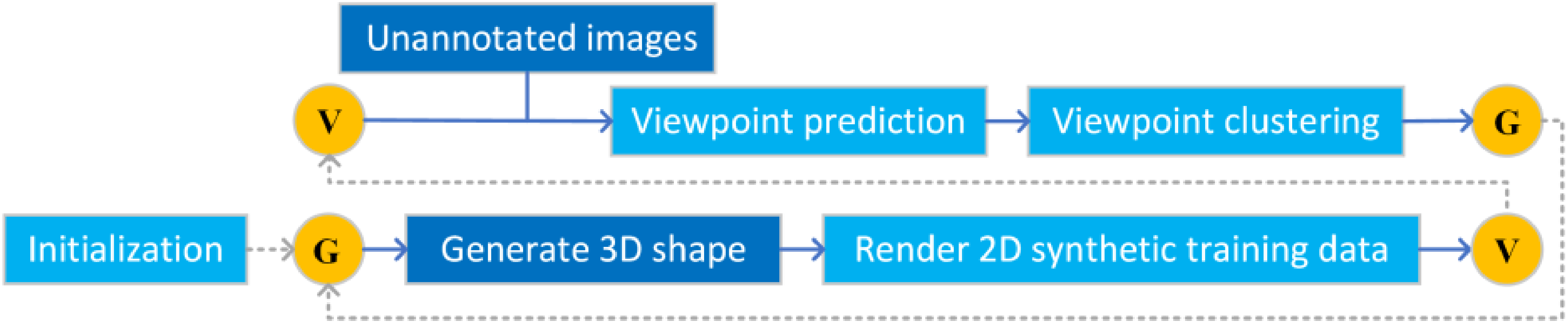
Our Solution

- A multiple projection GAN (MP)
  - One generator of 3D shapes
  - A projection layer
  - A set of discriminators, each for images of similar views
- A view prediction network (VP)
  - Predicting view information of images



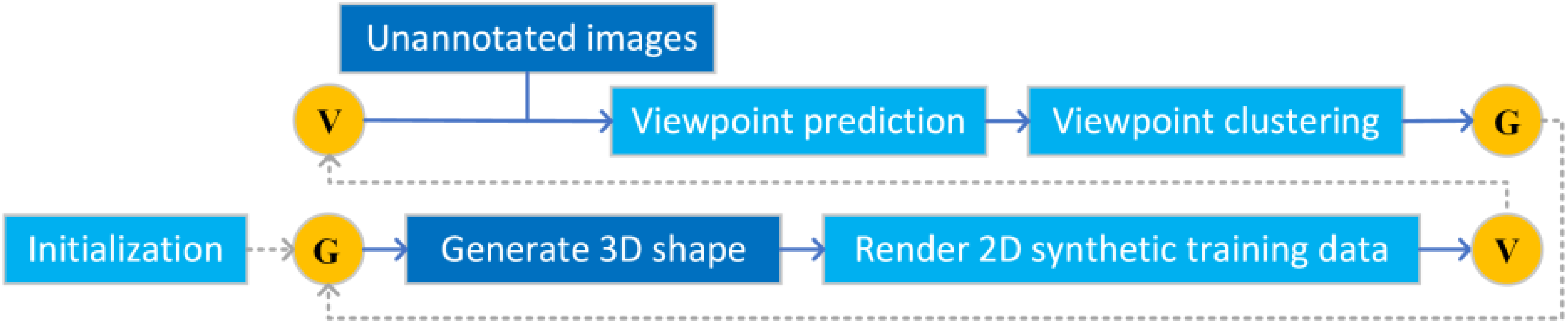
Our Solution

- Training two networks iteratively
  - Using rendering of 3D shapes generated by multi-projection GAN for VP training

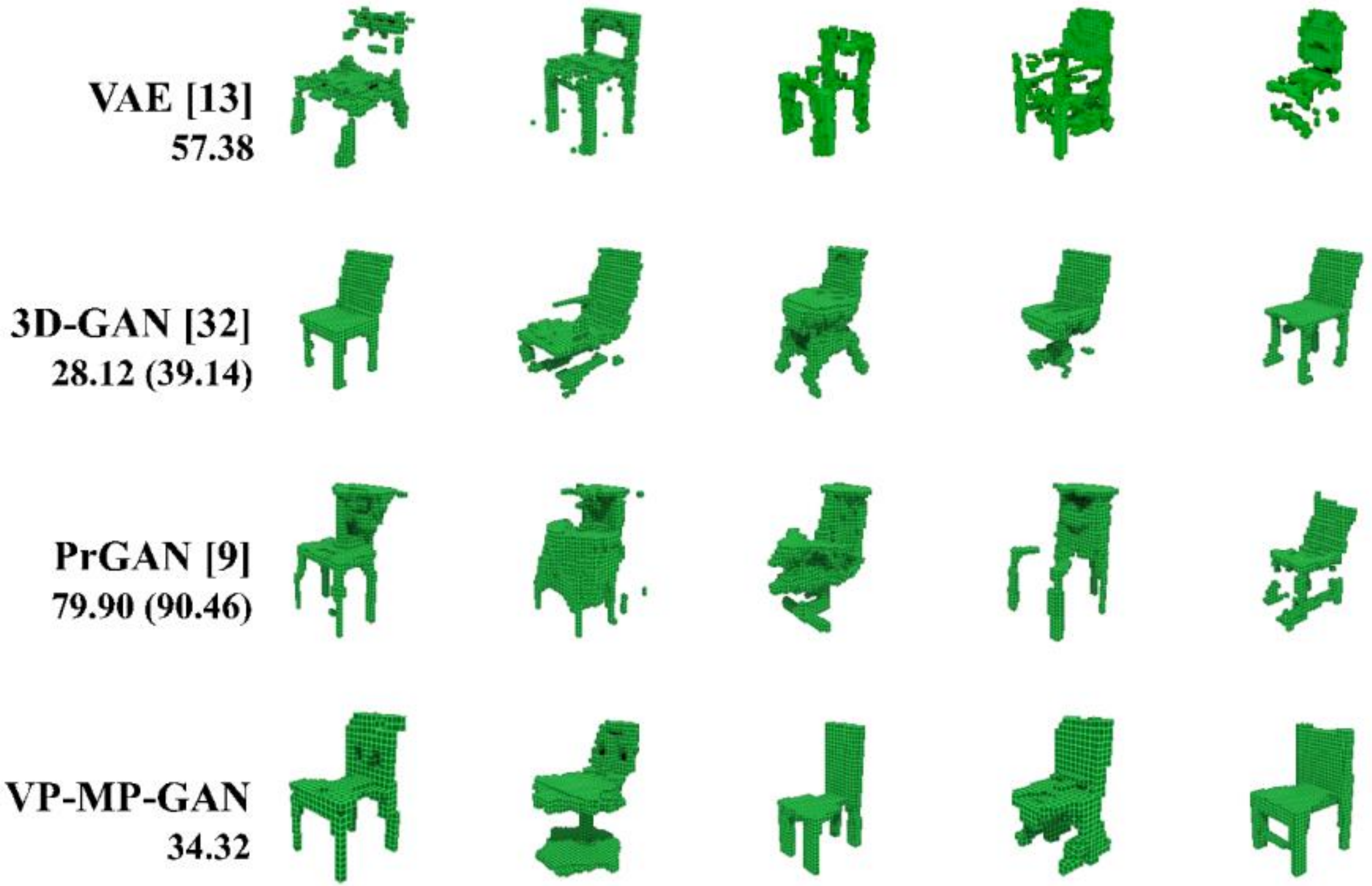


Our Solution

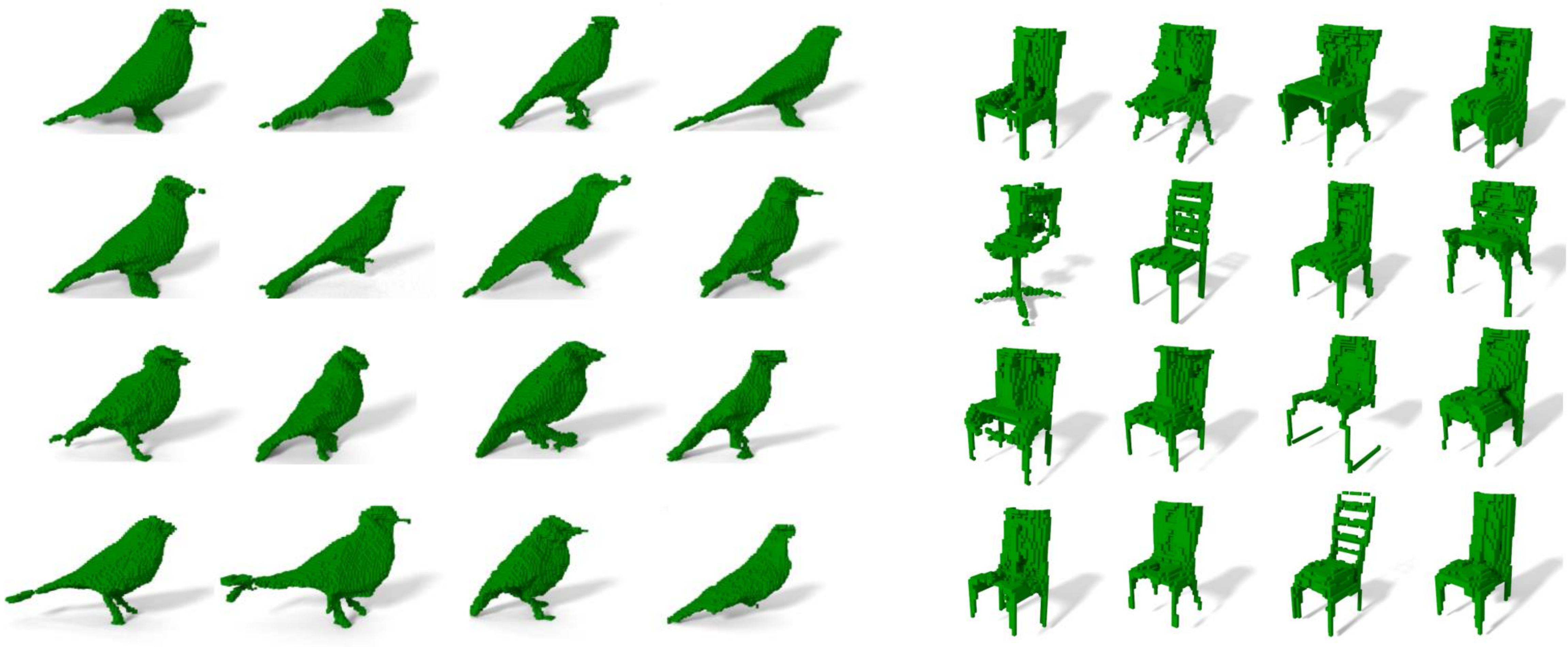
- Training two networks iteratively
  - Using rendering of 3D shapes generated by multi-projection GAN for VP training
  - Using VP to classify images according to their views
  - Training GAN with more discriminators, each corresponds to images in one class



Comparison

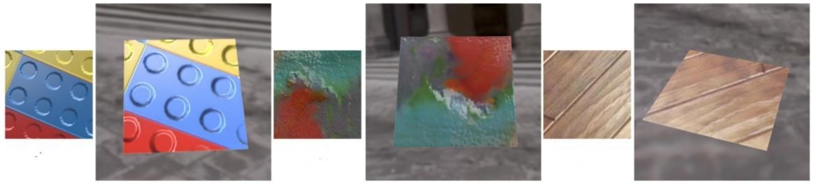


Results



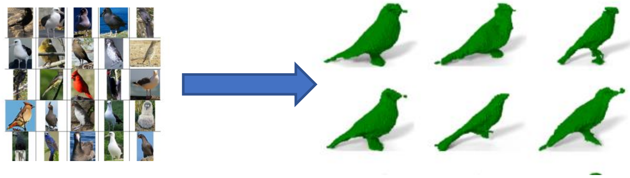
Our Efforts

Small labeled dataset



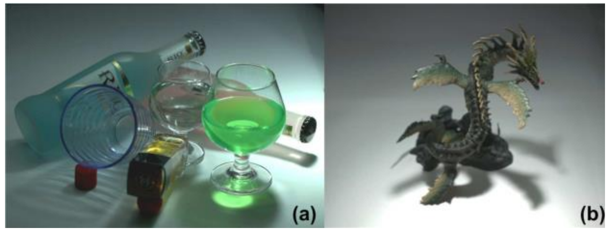
Self-augmented CNN training for SVBRDF modeling [SIGGRAPH 2017]

Dimensionality gap



Multi-projection GAN [CVPR 2019]

Variant representations



Kernel Nystorm for Relighting [SIGGRAPH 2009]

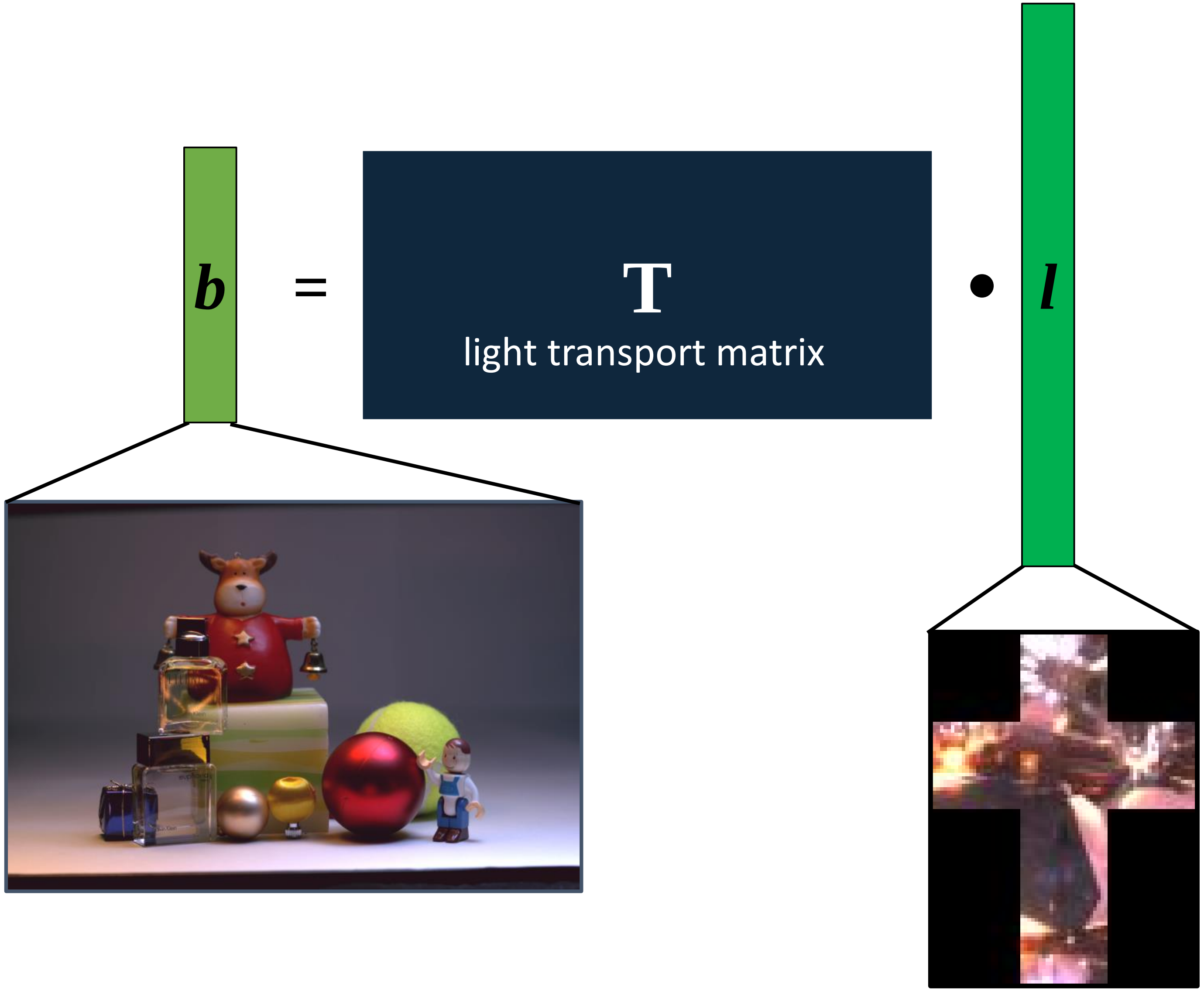
Image based Relighting [SIGGRAPH 2015]

Our Goal

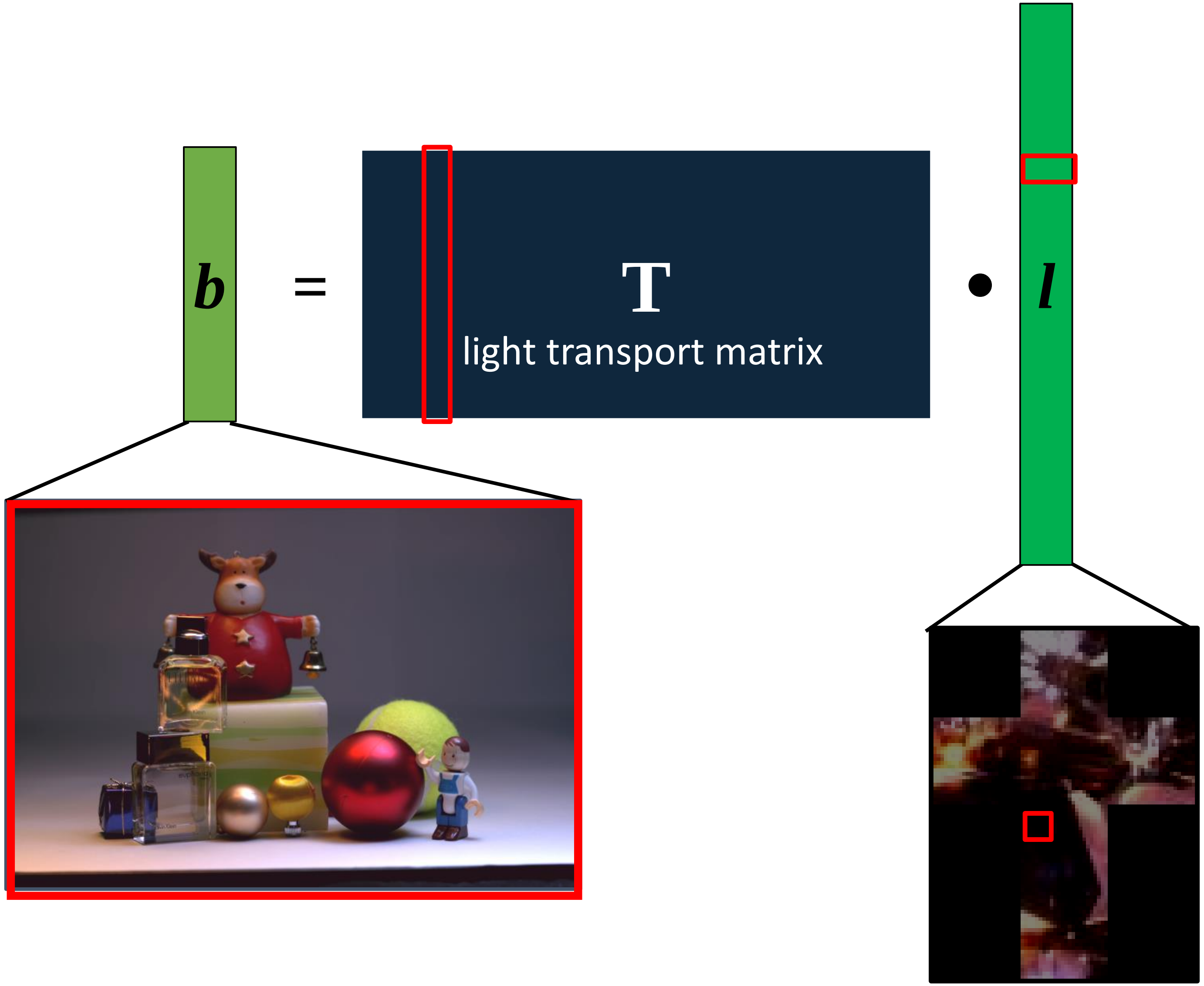
- Relighting a real-world scene with a small set of images
  - Fixed view, arbitrary lighting
  - No scene geometry and material information

Jiaping Wang, Yue Dong, Xin Tong, Zhouchen Lin, Baining Guo, Kernel Nytrom for Light Transport, ACM SIGGRAPH 2009

Light Transport Matrix for IBL

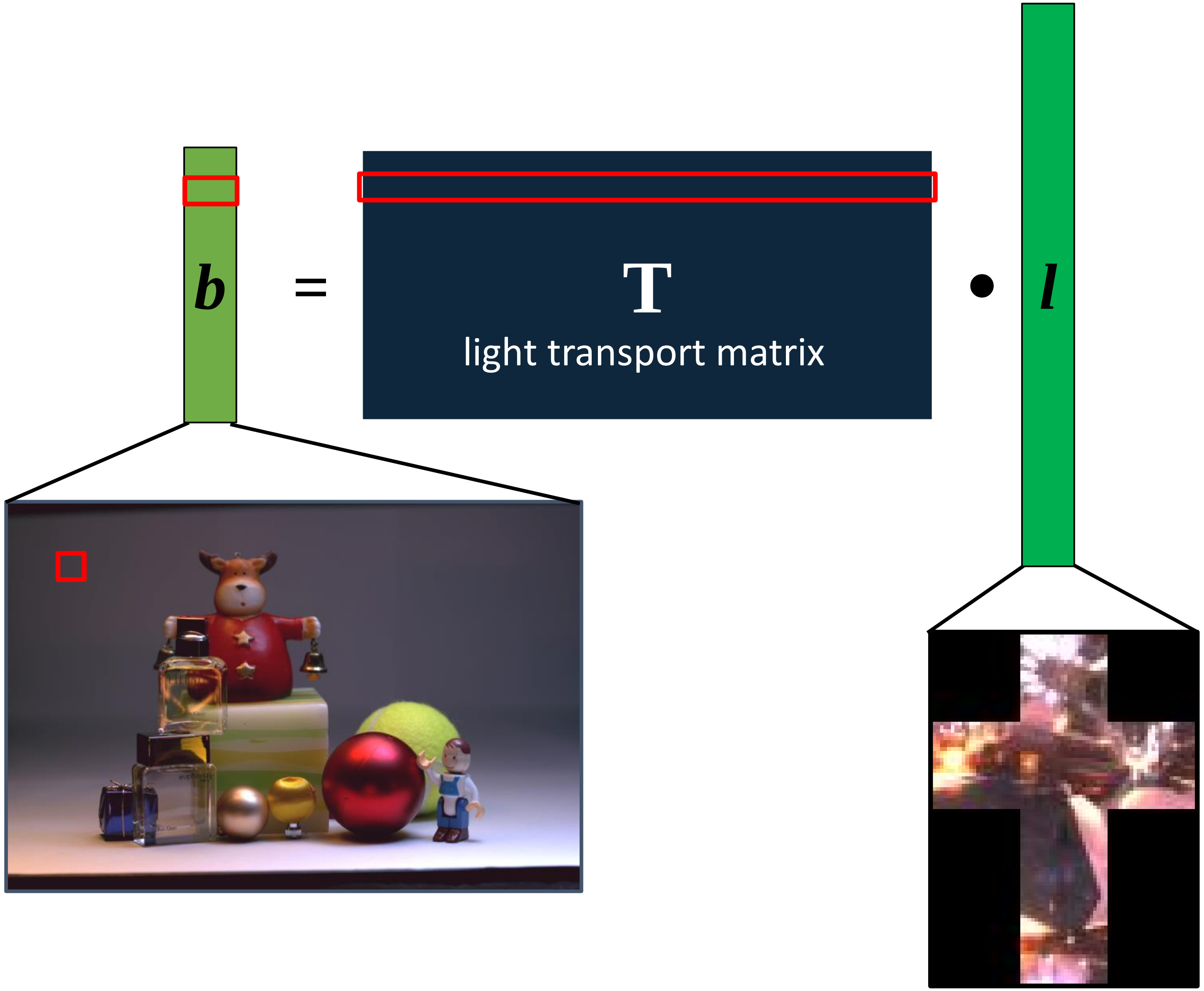


Light Transport Matrix for IBL



Microsoft Research 微软亚洲研究院

Light Transport Matrix for IBL



Key Challenge

- Reconstructing light transport matrix from few images
  - Avoid directly sampling the light transport matrix with dense images

Our Key Observation

- The light transport matrix of a scene is always low rank
  - With proper non-linear map, the rank could be lower due to linear & non-linear coherence inside the matrix

Our Solution

- Nyström scheme for low-rank matrix approximation

$$T = \begin{bmatrix} \mathbf{A} & \mathbf{R} \\ \mathbf{C} & \mathbf{B} = ? \end{bmatrix}$$

$$\mathbf{B} \approx \mathbf{C} \cdot \mathbf{A}^{+} \cdot \mathbf{R}$$

$\square^{+}$ denotes pseudo-inverse

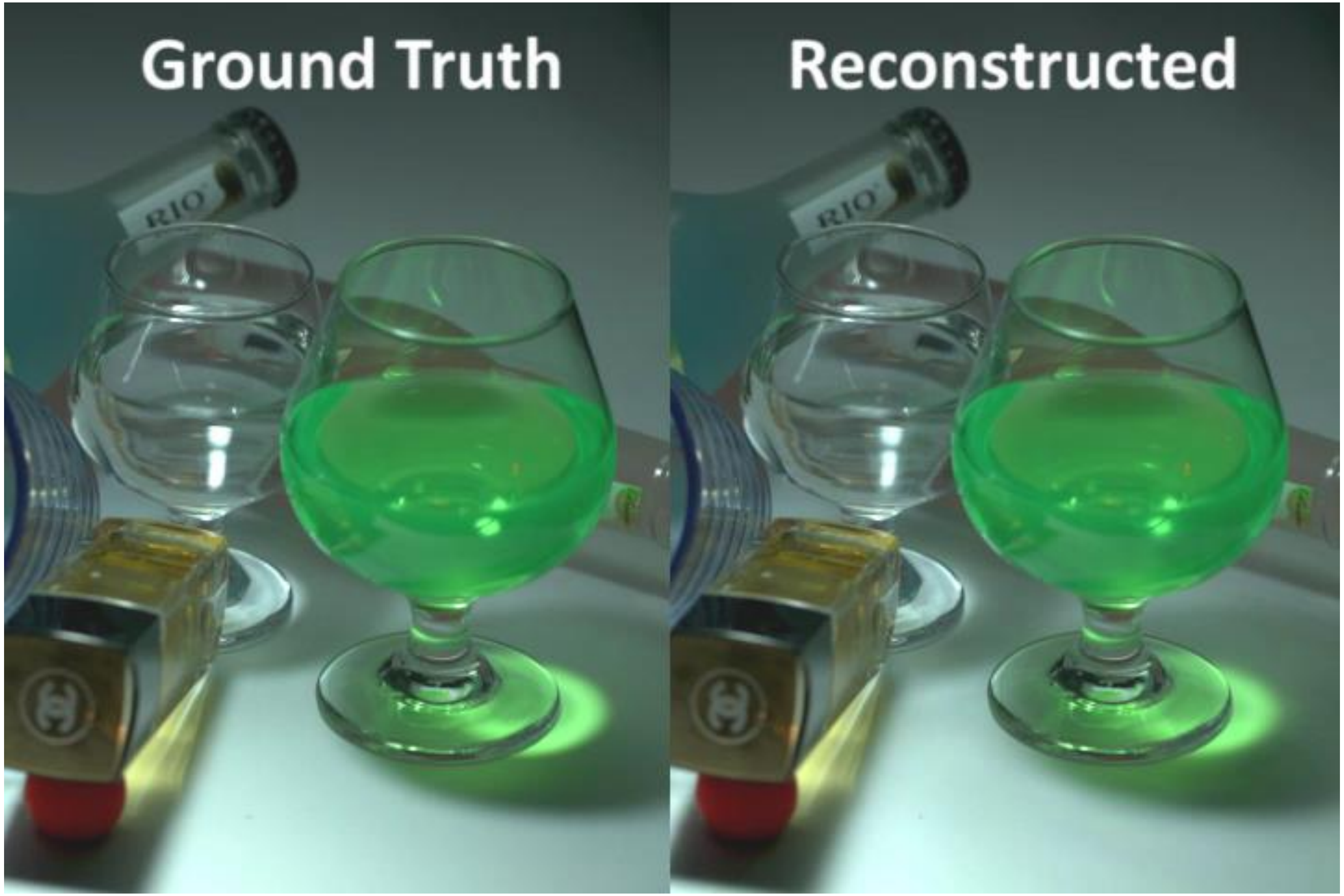
Our Solution

- Kernel Nyström scheme for low-rank matrix approximation
  - An inversible kernel map can further reduce the rank of light transport matrix
  - Optimize the kernel by minimizing nuclear rank of the kernel matrix

$$f(\mathrm{T}) = \begin{bmatrix} f(\mathrm{A}) & f(\mathrm{R}) \\ f(\mathrm{C}) & f(\mathrm{B}) = ? \end{bmatrix}$$

$$f(\mathrm{B}) \approx f(\mathrm{C}) \cdot f(\mathrm{A})^{+} \cdot f(\mathrm{R})$$

Results

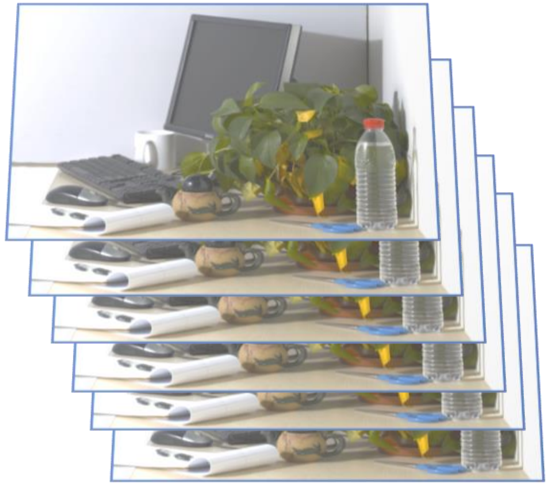


Our Goal

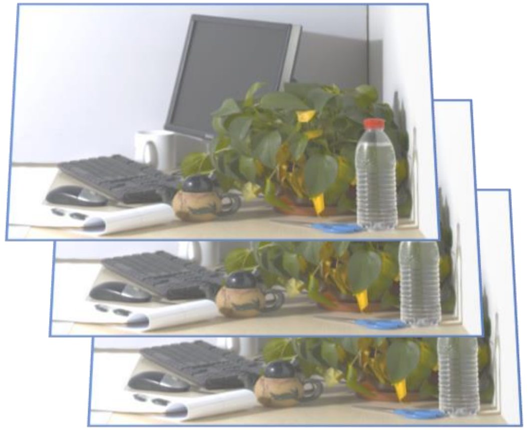
- An efficient image based relighting solution
    - As simple as possible device setup

Our Goal

- An efficient image based relighting solution
  - As simple as possible device setup
  - As few as possible images



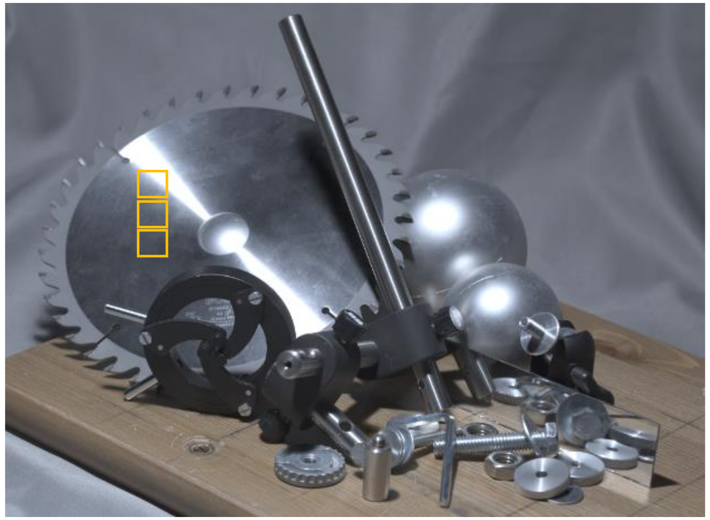
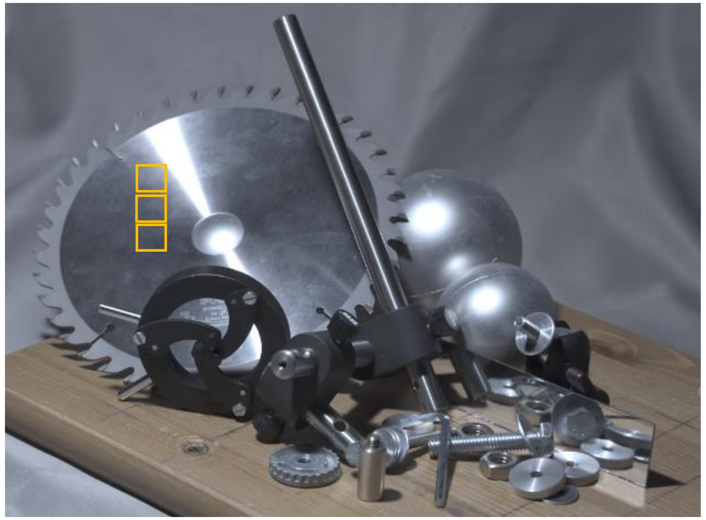
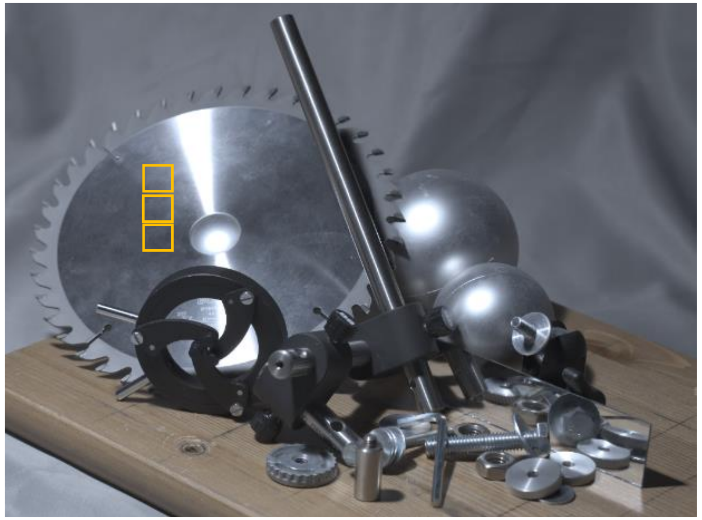
Required images
by previous methods



Required images
by our method

Our Key Observations

- Local non-linear coherence
  - Among nearby pixels
  - Among nearby lightings

Our Key Idea

- Model the relighting as a regression problem
  - Using neural networks for modeling relighting effects of local pixels (each light transport matrix element)
  - Each element as function of pixel position and lighting direction/position
  - Different image region with different neural networks
  - Leverage the non-linear coherence among all elements

$$I(\mathbf{p}, \mathbf{l}) = \Phi(\mathbf{p}, \mathbf{l})$$

Our Key Idea

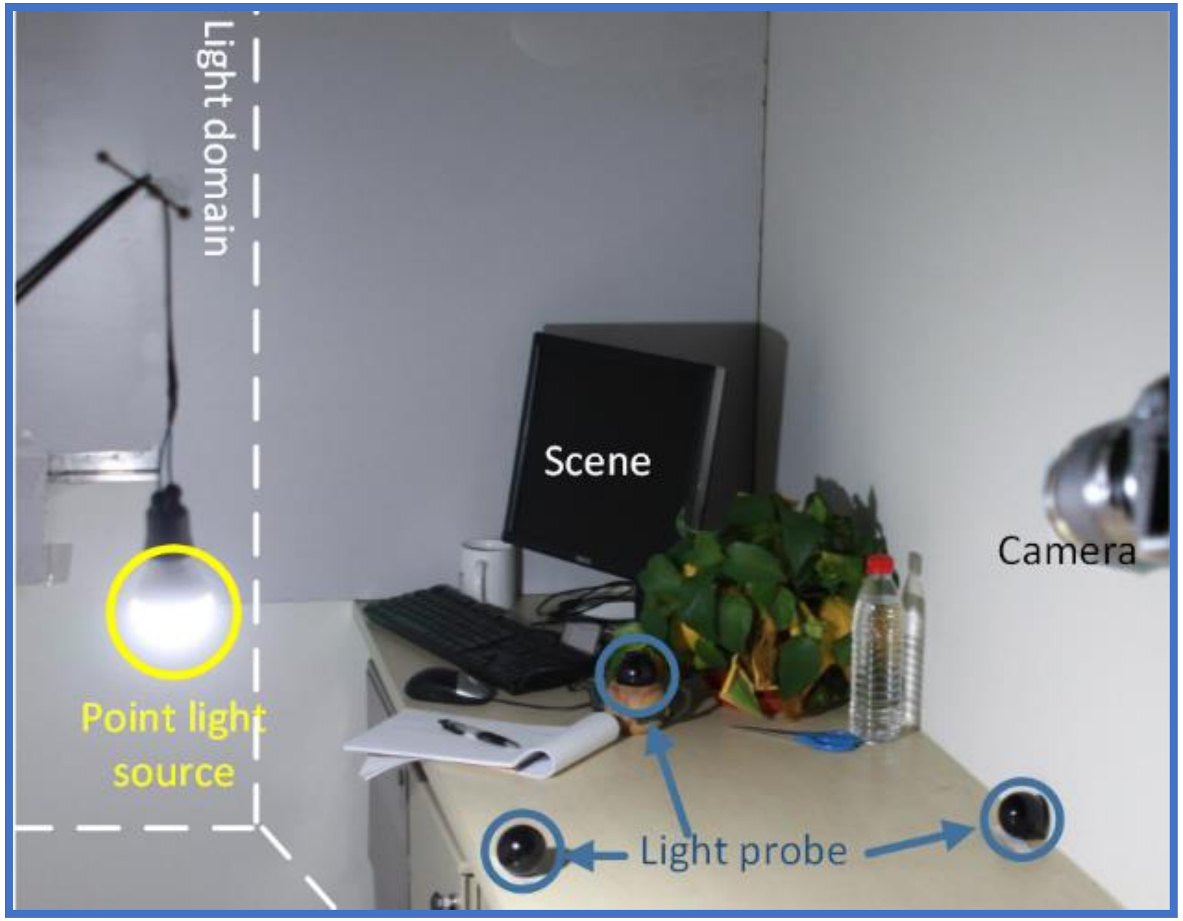
- Model the relighting as a regression problem
  - Regress the neural networks with pre-captured images under random lighting
  - Predict the relighting effects with result neural networks

Captured images → Neural network training → Relighting

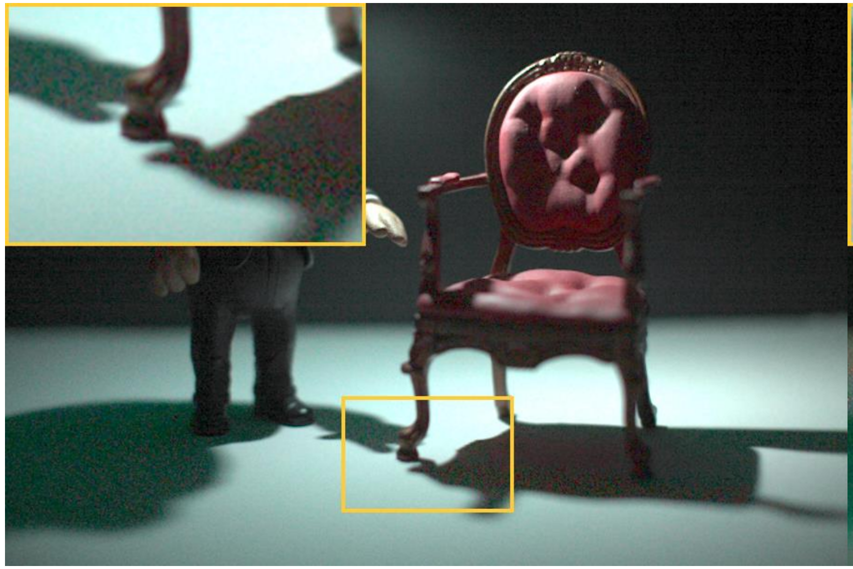
Offline training stage | Online relighting stage

Our Solution

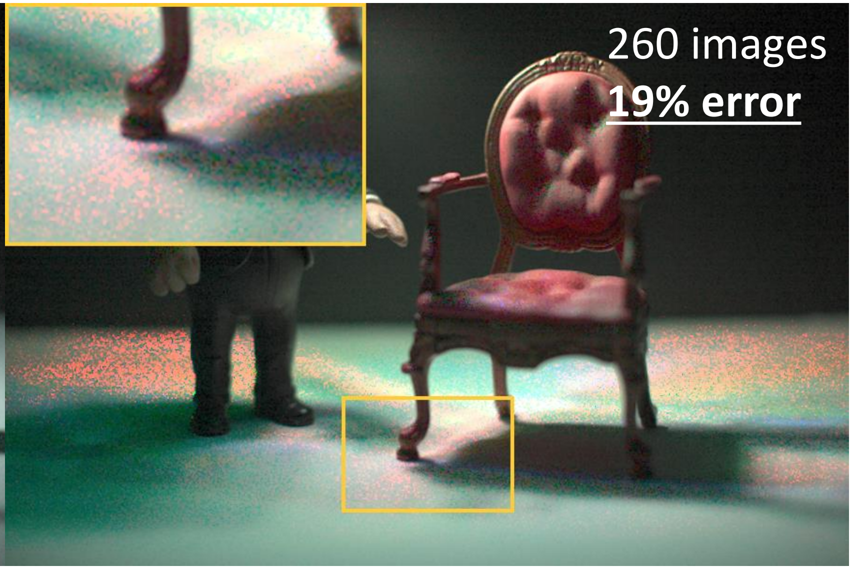
- Device setup
  - Hand moved point light source
  - Known 3D lighting position
  - Fixed view for image capturing



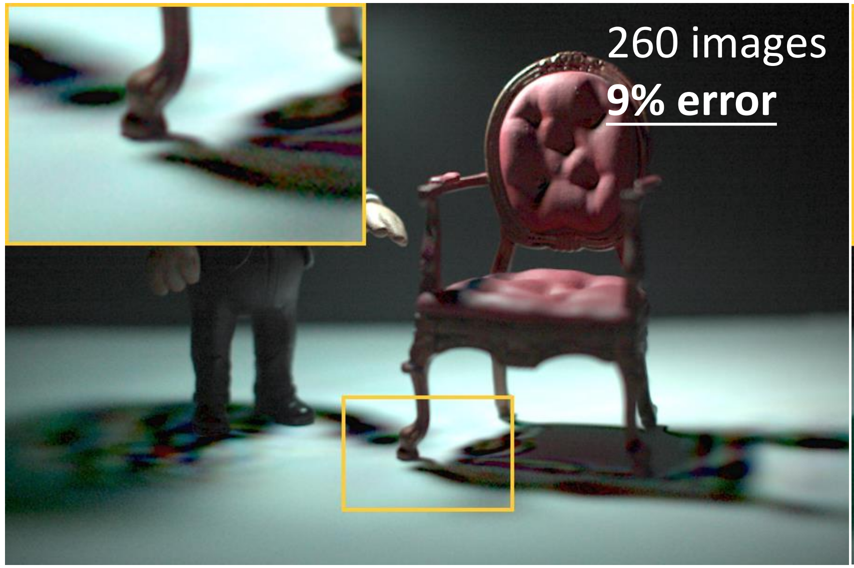
Results



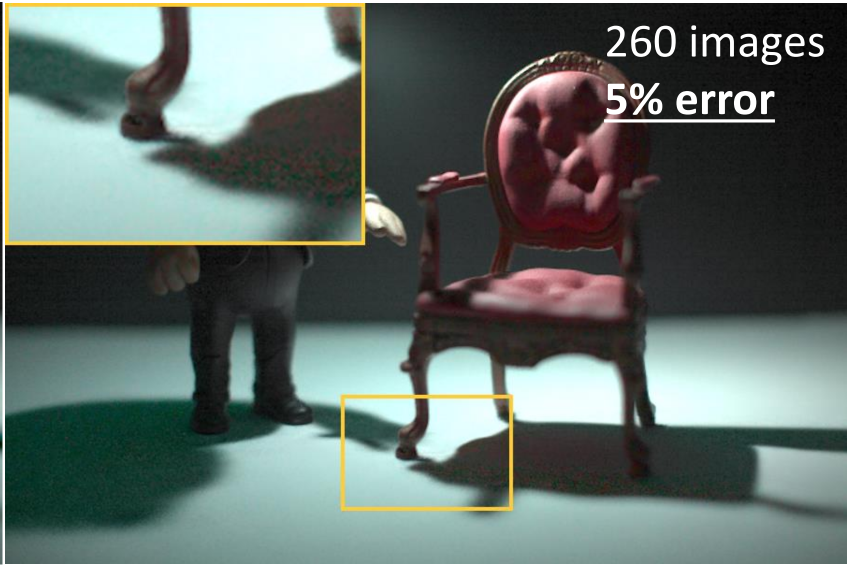
Captured photo



[Wang et al. 2009]



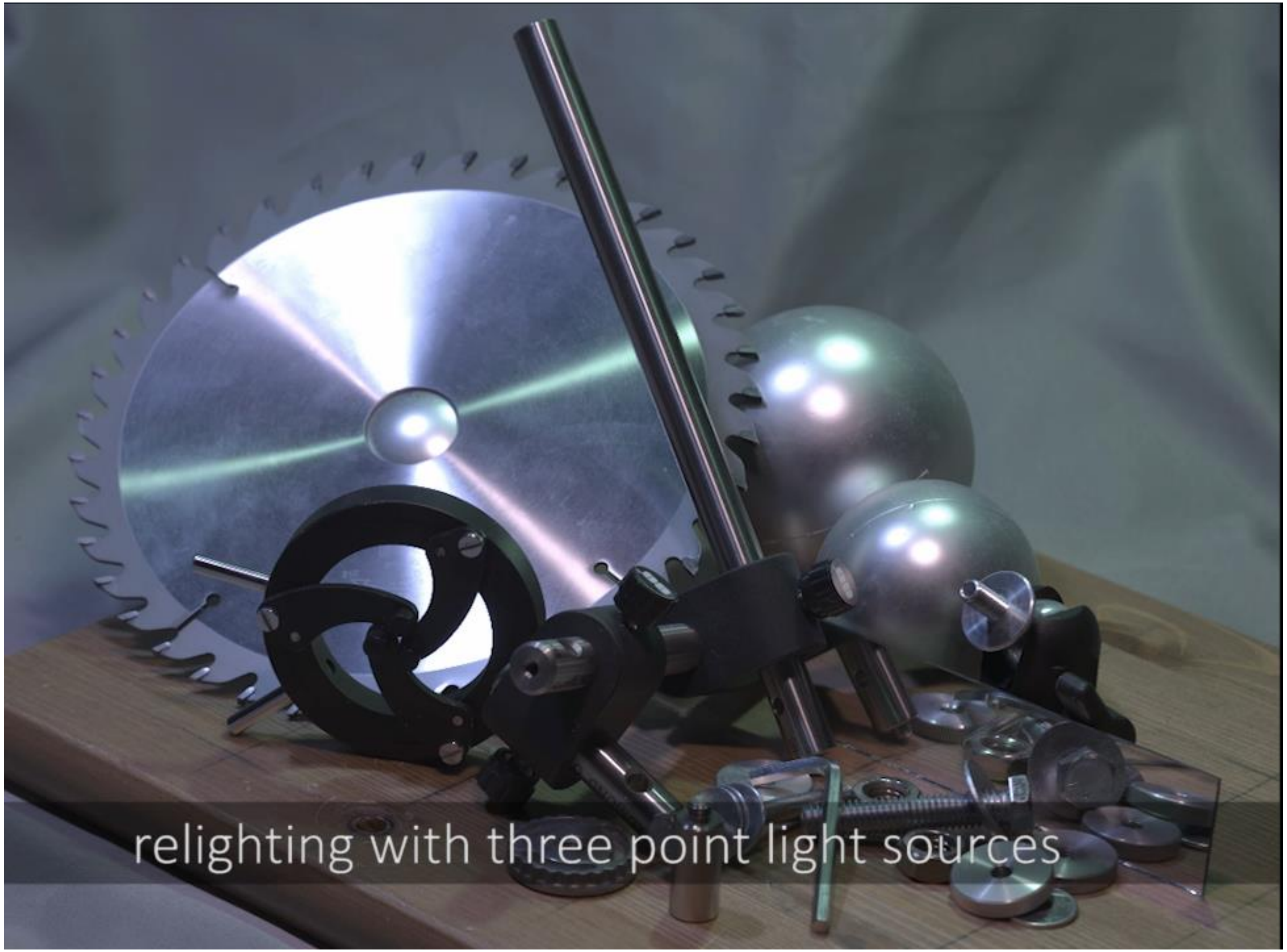
[O’Toole et al. 2010]



Our method

Results

Results



What I Learned from This Journey?

- Data driven ≠ deep learning
  - Try simple method first, especially when your data is not large
- Domain knowledge is critical
  - Reduce the data required for training
  - Decompose the problem into simpler one
  - Make the model robust and easy to train
- Problem formulation is important
  - New formulation leads to new solutions

Outline

- An overview of data driven graphics
- Key challenges in data driven graphics and our exploration
- Future directions

Challenges: Data

- Developing automatic and fast capturing systems
- High quality and easy-to-use modeling tools for end-users
  - With the help of sparse sketches or image/video

Challenges: Models

- General CNN models for 3D shape/material/motion analysis
- Powerful GAN model for high quality 3D data generation

Challenges: Learning Methods

- Integrating with the physical priors and constrains
- Bridging the gap between image/videos and high dimensional graphics data
- Allowing user control/editing

From Simple Tools to Intelligent Assistants

- Realizing the user intentions and design goals
- Converting abstract input (text/speech) into concrete 3D contents

Thanks