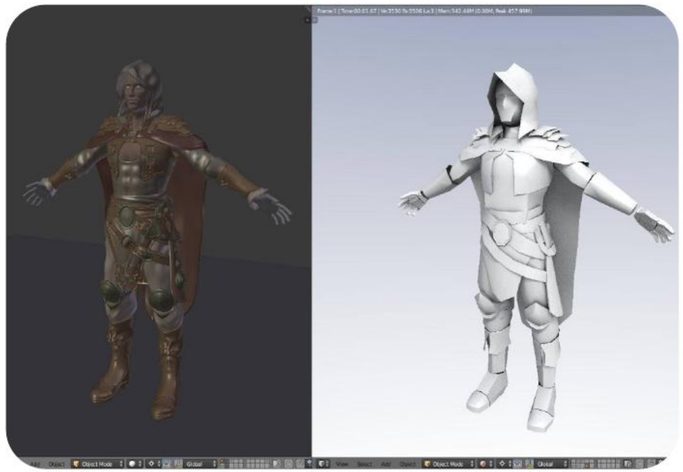
Deep Reflectance Volumes: Relightable Reconstructions from Multi-View Photometric Images

Sai Bi[1] Zexiang Xu[1,2] Kalyan Sunkavalli[2] Miloš Hašan[2] Yannick Hold-Geoffroy[2] David Kriegman[1] Ravi Ramamoorthi[1]

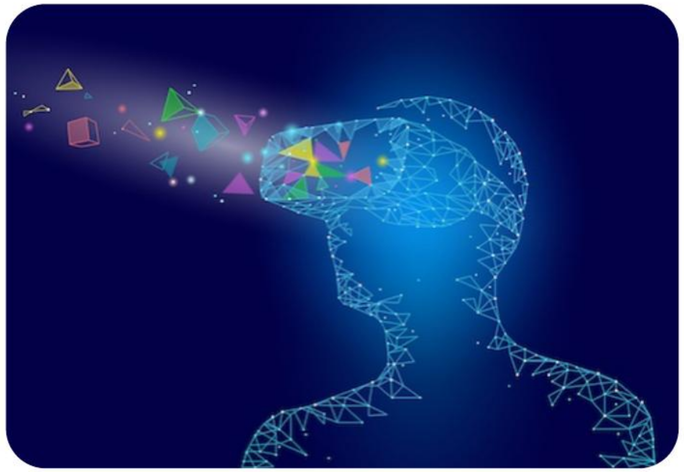
[1] UC San Diego [2] Adobe Research

Appearance acquisition

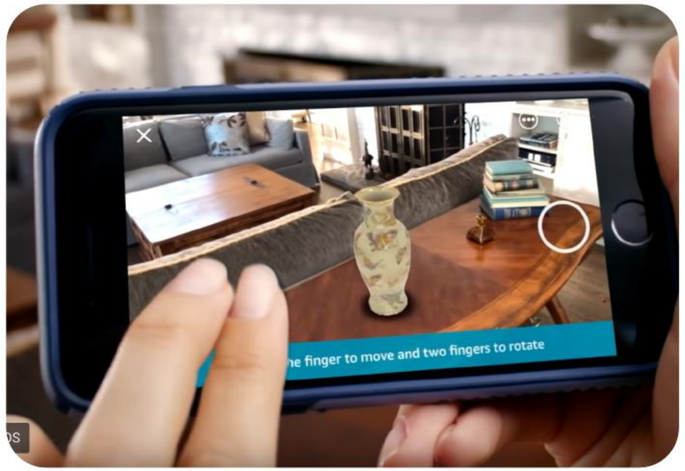


Video games

Cultural heritage preservation



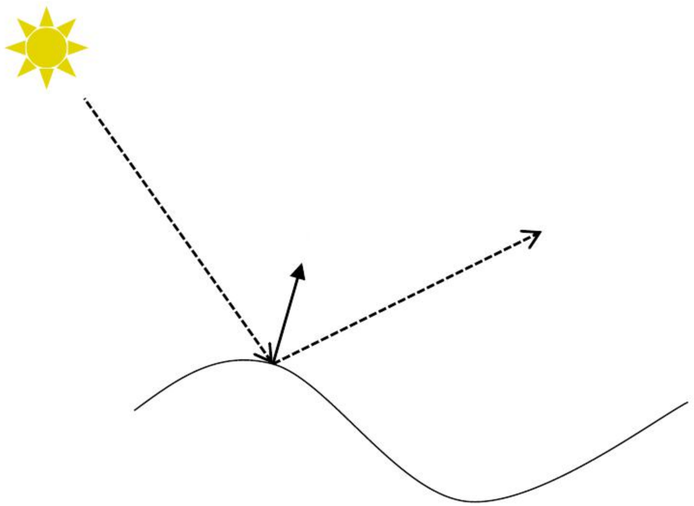
Virtual reality



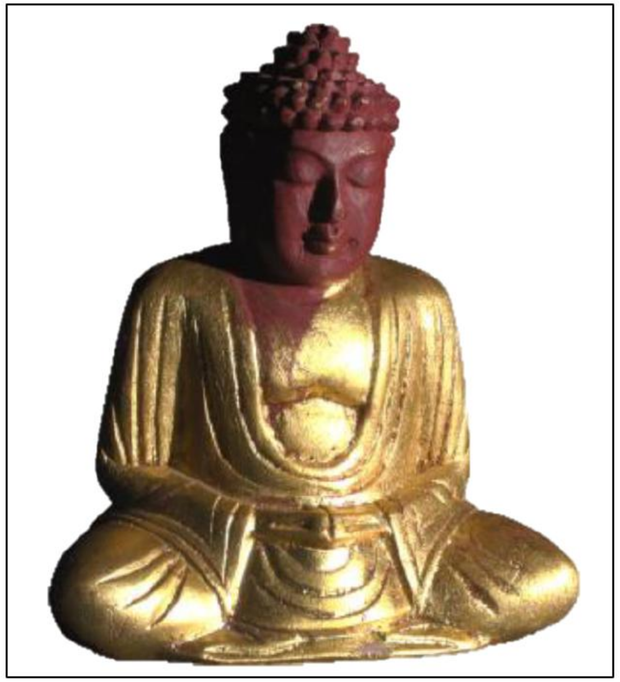
Augmented reality

Mesh-based appearance acquisition

Geometry



BRDF



Rendering

Mesh-based appearance acquisition

- Complex acquisition setups
  - High-end 3D scanners, light stages
- Multi-view stereo
  - Difficult to handle complex geometries

[Zhou et al. 2016]

[Xia et al. 2016]

[Kang et al. 2019]

[Nam et al. 2018]

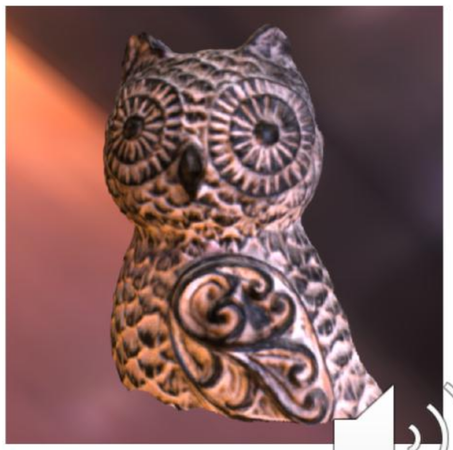
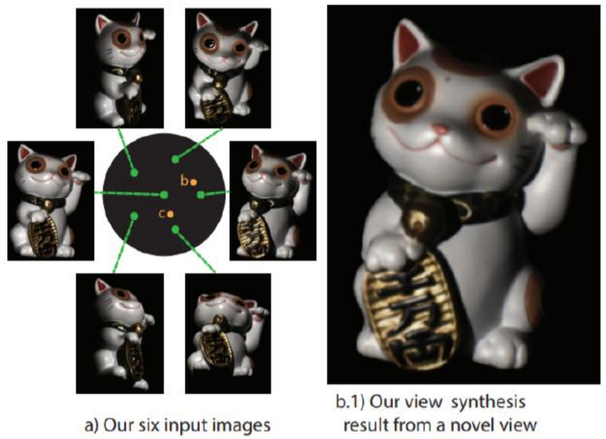

[Bi et al. 2020]

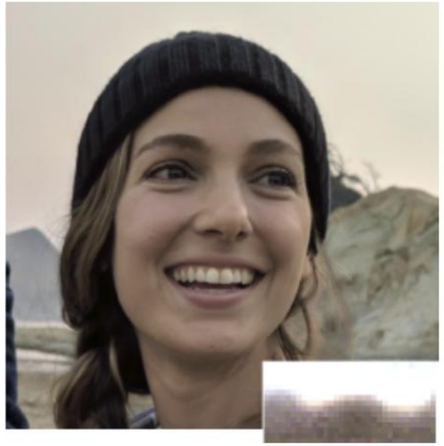
Image-based appearance acquisition

- Limited to a single task
- Need to capture images under different lighting conditions

[Xu et al. 2018]



[Xu et al. 2019]

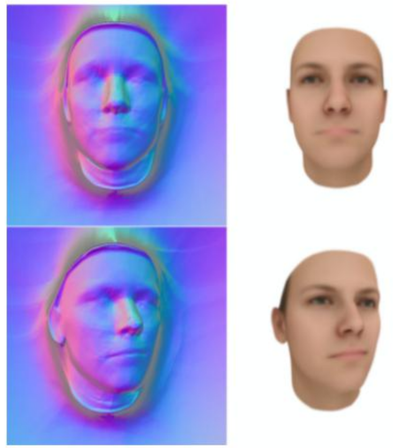


[Sun et al. 2019]

Volumetric representations

- Limited to view synthesis under fixed lighting conditions

[Sitzmann et al. 2019]



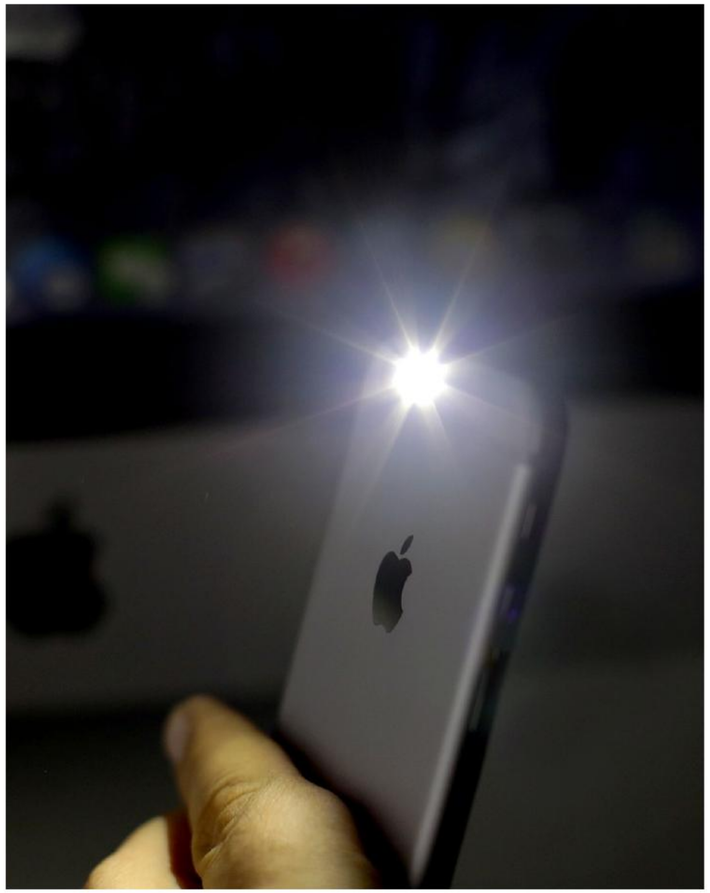
[Sitzmann et al. 2019]

[Lombardi et al. 2019]

[Mildenhall et al. 2020]

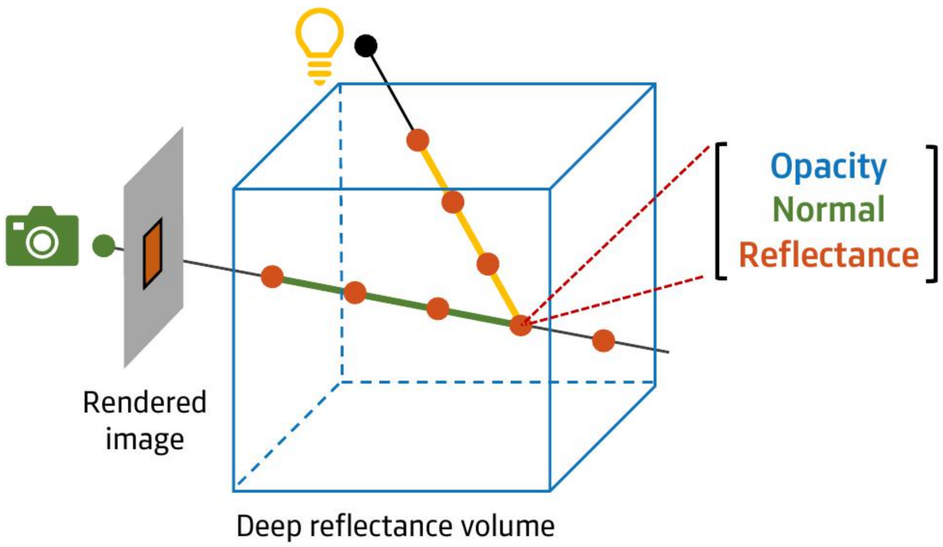
Joint view synthesis and relighting

Mobile phone captures with flashlight



Deep reflectance volumes

- Physically-based volumetric representation
- Differentiable raymarching framework

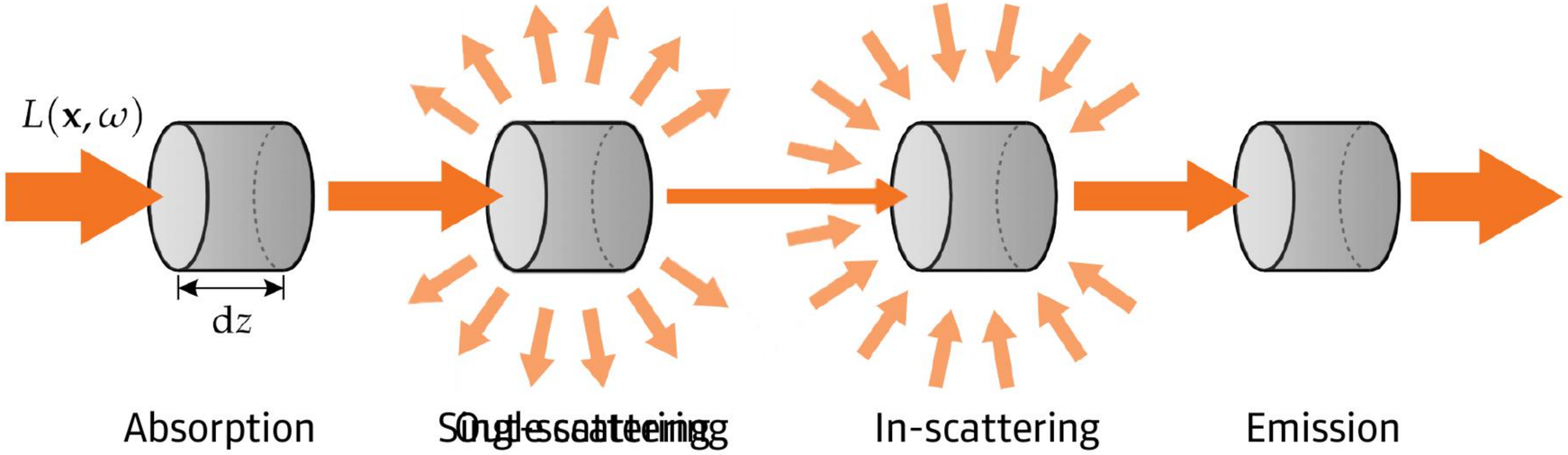


Overview

- Render with differentiable raymarching
- Learn the deep reflectance volumes
- Experiment results

Volume rendering

- Single point light
- Single scattering
- No emission

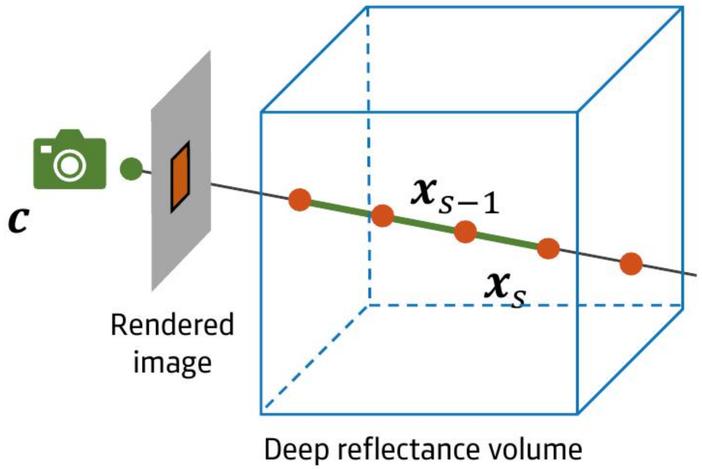


Discretized volume rendering

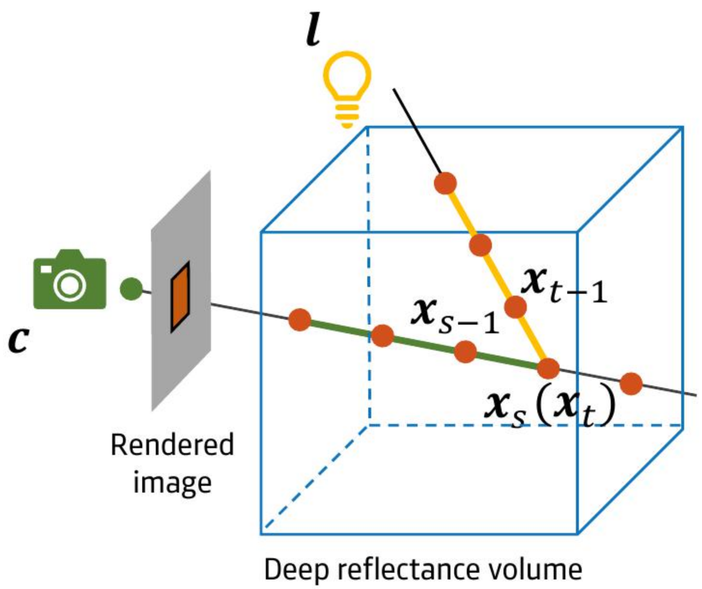
- Differentiable ray marching

$$\boldsymbol{x}_s = \boldsymbol{x}_{s-1} - \boldsymbol{\omega}_o \, \Delta x$$

View direction — Step size



Discretized volume rendering



- Opacity compositing

Light ray $\alpha_{l\to t} = \boxed{\alpha_{l\to(t-1)}} + \boxed{\left[1-\alpha_{l\to(t-1)}\right]\alpha(\boldsymbol{x}_t)}$

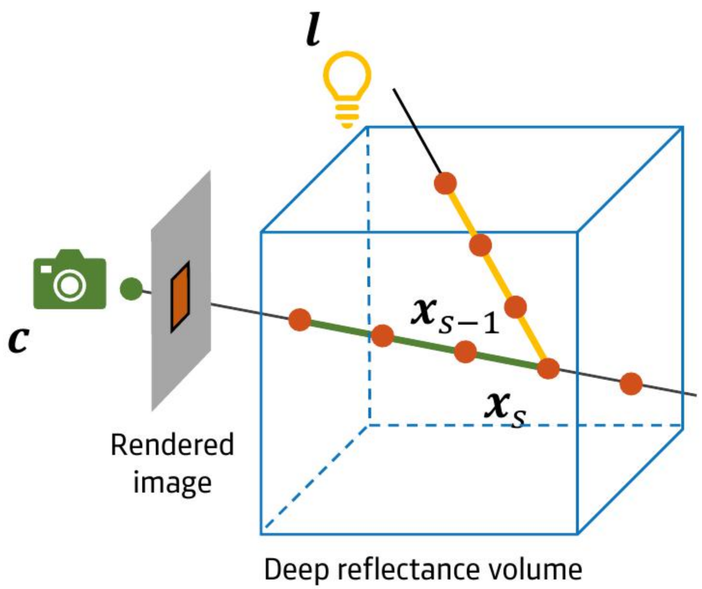
Accumulated opacity from $\boldsymbol{l}$ to $\boldsymbol{x}_t$

Opacity at $\boldsymbol{x}_t$

Identical during training

Camera ray $\alpha_{c\to s} = \alpha_{c\to(s-1)} + \left[1-\alpha_{c\to(s-1)}\right]\alpha(\boldsymbol{x}_s)$

Discretized volume rendering



- Reflected radiance at x_s

$$L(\boldsymbol{x}_s) = L_l(\boldsymbol{x}_s, \boldsymbol{\omega}_i) f_r(\boldsymbol{\omega}_o, \boldsymbol{\omega}_i, \boldsymbol{n}(\boldsymbol{x}_s), R(\boldsymbol{x}_s))$$

Lighting intensity

normal

Reflectance

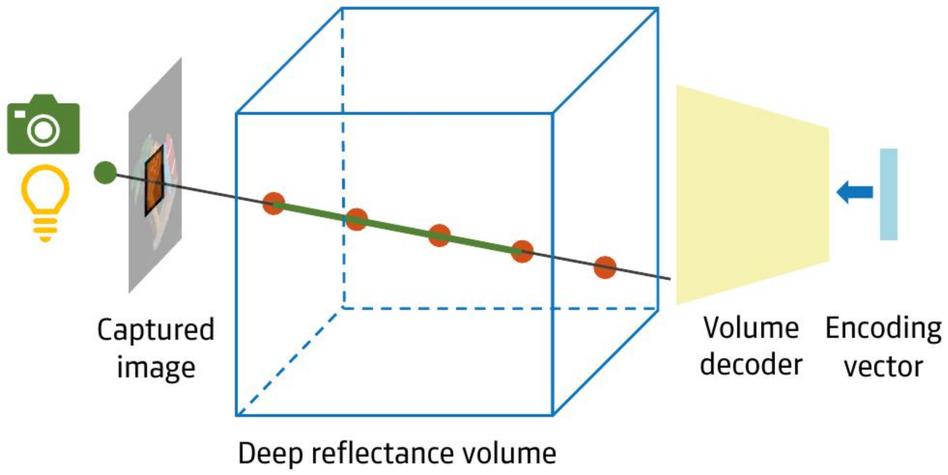
Diffuse albedo

Specular roughness

$$L_s = L_{s-1} + \left(1 - \alpha_{c\to(s-1)}\right)\left(1 - \alpha_{l\to(t-1)}\right)\alpha(\boldsymbol{x}_s)\, L(\boldsymbol{x}_s)$$

Learning deep reflectance volumes

- Start from a encoding vector randomly initialized with Gaussian
- Decode into an 8-channel $128 \times 128 \times 128$ volume
- Train on images under collocated camera and light
- Generalize to arbitrary novel view and light



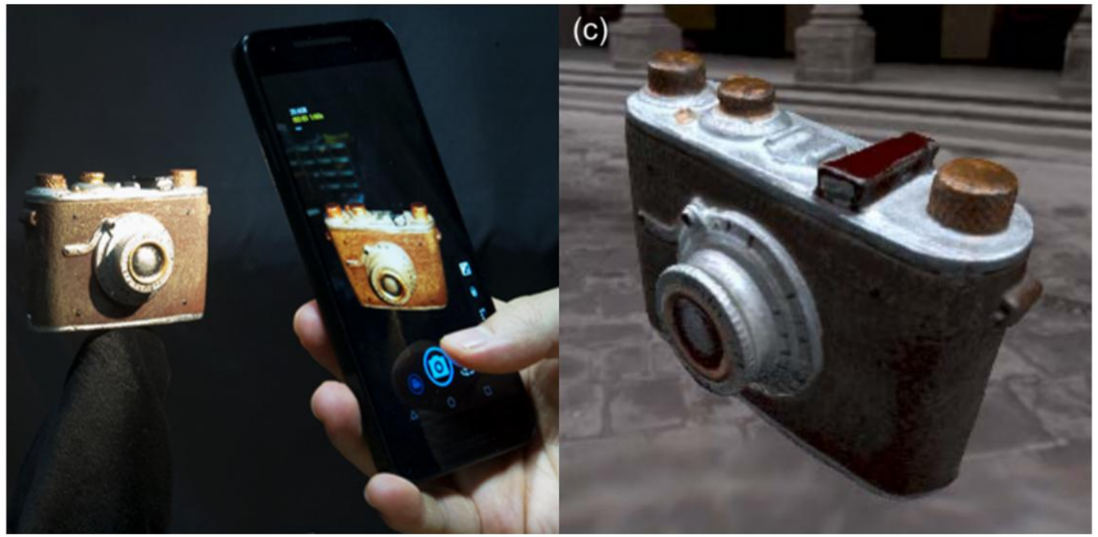
Loss functions

- Image reconstruction loss
  - Minimize difference between predicted pixel colors and ground truth
- Accumulated opacity loss
  - Most camera rays hit either the object or hit the background
- Opacity sparsity loss
  - Enforce the opacity to be spatially sparse

$$\sum_k \left|\left|L_k - \tilde{L}_k\right|\right|_2^2 + \beta_1 \sum_k [\log(\alpha_{c_k \to s'}) + \log(1 - \alpha_{c_k \to s'})] + \beta_2 \sum_x ||\nabla_x \log \alpha(x)||$$

Experiment results

- Compare to state-of-the-art mesh-based methods [Nam et al. 2018]
  - Use multi-view stereo to reconstruct geometry
  - Solve for per-vertex BRDFs via optimization



[Nam et al. 2018]

Comparison to mesh-based methods

PONY: sample input images

Comparison to mesh-based methods

HOUSE: sample input images

Additional results

CARTOON: our renderings

Additional results

ANIMALS: our renderings

Comparison on synthetic data

STATUE: sample input images

Environment map rendering

Ours

Ground truth

Material editing

Before editing

After editing

Summary

- Deep reflectance volumes that model both geometry and reflectance
- A physically-based differentiable rendering module
- Support joint view synthesis and relighting

Future works

- Sparse volumetric representations
- Implicit volumetric representations using MLPs
- Extension to other materials like hair, fur, or glass

Thank you!

Acknowledgements

- ONR grants N000141712687, N000141912293, N000142012529
- NSF grant 1617234
- Adobe
- Ronald L. Graham Chair
- UC San Diego Center for Visual Computing