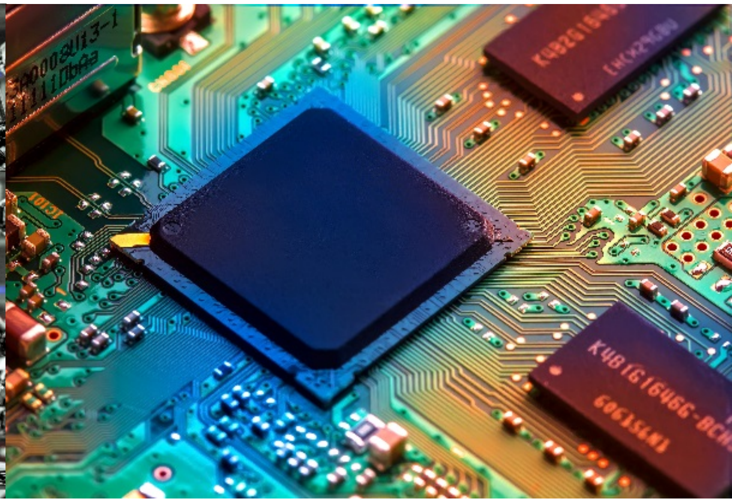
Scene Reconstruction From Monocular Image and Photometric Stereo

Dohun Jeong

Advised by Prof. Minh N. Do



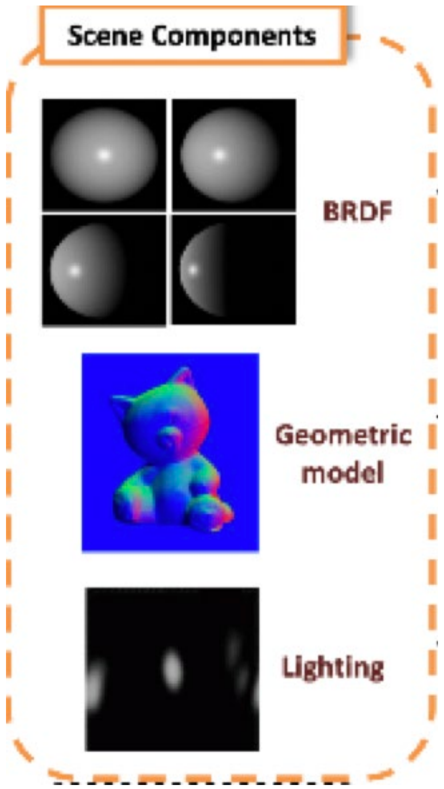
ILLINOIS

Electrical & Computer Engineering

GRAINGER COLLEGE OF ENGINEERING

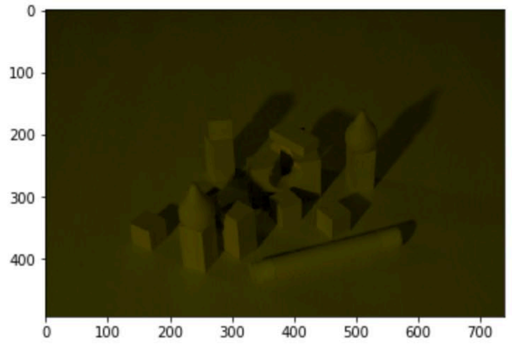
Introduction

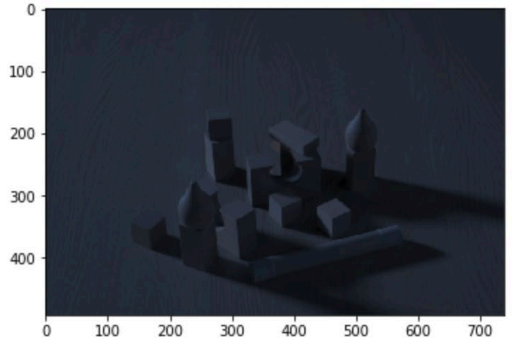
- Motivation: Inferring scene information from photograph
  - Scene editing in augmented reality (AR) applications
  - Intrinsic Image Decomposition in image processing
- Goals in this thesis
  - BRDF recovery
  - Photorealistic style transfer
  - Depth recovery

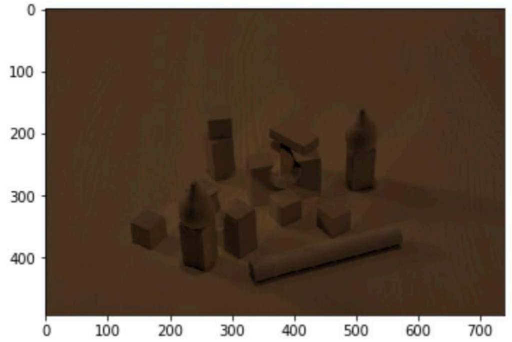


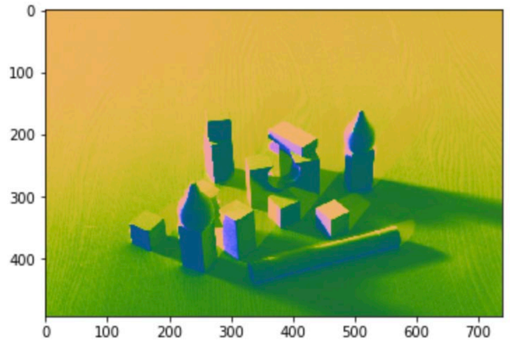
Introduction - Background

- Classical computer vision algorithms
  - Known light sources
  - Lambertian shading
  - No shadows
  - No interreflections









Introduction – Recent Works

- Image synthesis perspective
  - Rendering equation (Kajiya 1986)

$$L_o(\mathbf{x}, \omega_o, \lambda) = L_e(\mathbf{x}, \omega_o, \lambda) + \int_\Omega f_r(\mathbf{x}, \omega_i, \omega_o, \lambda) L_i(x, \omega_i, \lambda)(\omega_i \cdot \mathbf{n}) d\omega_i$$

  - Monte Carlo path tracing (Veach 1997)
- Scene parameterization
  - Analytic BRDF
  - Triangular mesh
  - Mesh light source

Introduction – Recent Works

- Variational inference perspective (Kingma and Welling, 2013, Balakrishnan et al., 2019)

Methods

$$\underset{\Theta}{\operatorname{argmin}}\ E(\Theta) = \sum_{j}^{N} |I_C^j - \tilde{I}_R^j|_1$$

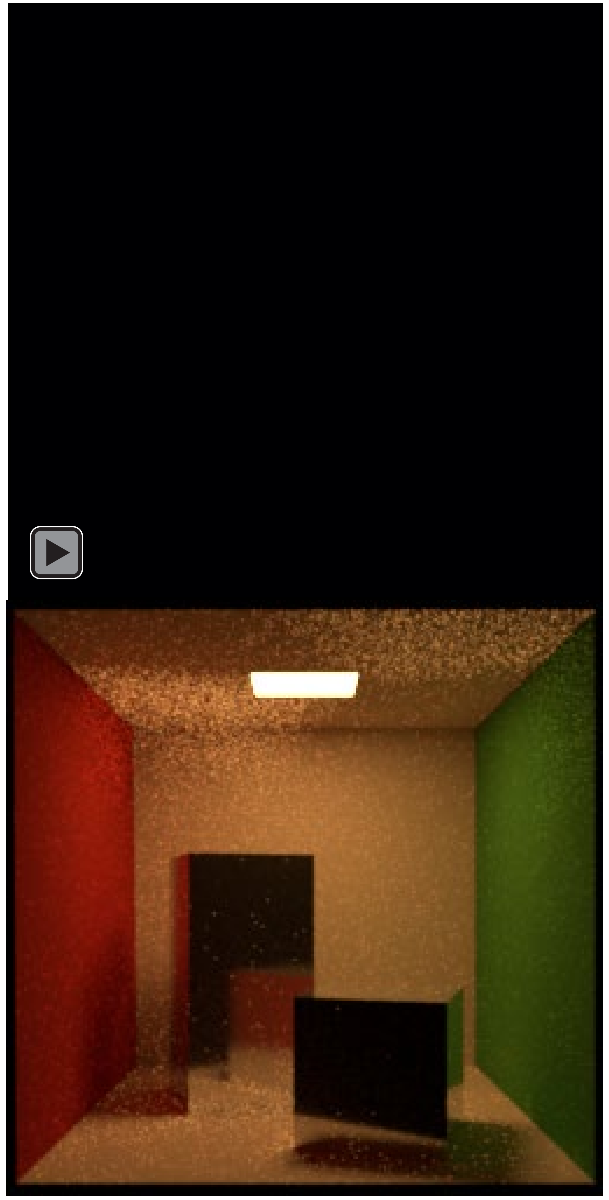
- Differentiable Path Tracing
  - Mitsuba 2 Inverse Renderer (Nimier-David, Vicini, et. al., 2019)
  - Analytic BRDF: Linear interpolation of diffuse and specular
    - GGX distribution for specular reflection, Lambertian for diffuse
  - Lighting:
    - BRDF recovery: Fixed point source or mesh light
    - Photorealistic Style Transfer: Environment Map and mesh light
- Hardware
  - NVIDIA RTX 2060 SUPER, 6GB GDDR

Methods

- Variational Inference
  - Objective: Find the most probable deprojected 3D volumetric data *y*, given a projected 2D image *x*, with latent space vector *z*
  - 3 function approximations using Convolutional Neural Net
    - Posterior network ~p(z|y)~Normal(mu_psi(y), sigma_psi(y))
    - Prior network p(z|x))~Normal(mu_phi(x), sigma_phi(x))
    - Deprojection network y = g(x, z)

BRDF Estimation

- Top: Initialized as diffuse surfaces
- Bottom: Target photo
- Fix lighting condition to a “pure-flash” configuration
  - – Known light source and location after camera calibration



Conclusion

- Scene geometry recovery remains the most challenging
  - Scene geometry also approximately differentiable through sampling (Loubet et al., 2019)
- Assuming scene geometry is known, a wide range of BRDF can be represented
  - Restrictions: Multi-lobed BRDF

References

- Kajiya, James T. "The rendering equation." *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*. 1986.
- Veach, Eric. *Robust Monte Carlo methods for light transport simulation*. Vol. 1610. PhD thesis: Stanford University, 1997.
- Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." *arXiv preprint arXiv:1312.6114* (2013).
- Balakrishnan, Guha, et al. "Visual Deprojection: Probabilistic Recovery of Collapsed Dimensions." *Proceedings of the IEEE International Conference on Computer Vision*. 2019.
- Loubet, Guillaume, Nicolas Holzschuch, and Wenzel Jakob. "Reparameterizing discontinuous integrands for differentiable rendering." *ACM Transactions on Graphics (TOG)* 38.6 (2019): 1-14.
- Nimier-David, Merlin, Vicini, Delio et al. "Mitsuba 2: a retargetable forward and inverse renderer." *ACM Transactions on Graphics (TOG)* 38.6 (2019): 203.