LIME: Live Intrinsic Material Estimation

Abhimitra Meka1,2  Maxim Maximov1,2  Michael Zollhöfer1,2,3  Avishek Chatterjee1,2  Hans-Peter Seidel1,2  Christian Richardt4  Christian Theobalt1,2

1Max-Planck-Institute for Informatics  2Saarland Informatics Campus  3Stanford University  4University of Bath

1. Goal & Motivation

- Estimating material (BRDF/reflectance) of general-shape objects from a single RGB image in real time
- AR applications impose new constraints on the problem of material estimation
  - Real-time
  - Uncontrolled lighting
  - In-the-wild setting
  - Limited sensor input
  - General shape objects

2. Model & Pipeline

- We decompose the input image into its physical constituents, inspired by the rendering equation:
  \[ L(x, \omega_i) = \int f(\alpha, \omega_i, \omega_i)(\alpha, \omega_i)E(\omega_i)d\omega_i \]
- Surface radiance (I) is given by the integral over the product of surface reflectance (f) and the incident illumination (E)
- The Blinn-Phong reflection model is used to simplify the rendering equation
- The incident lighting is parameterized by an environment map
- We use 5 CNN’s that are tailored to perform specific sub-tasks on the input image
  - All the networks are trained jointly in an end-to-end fashion by combining a ground-truth loss with a novel perceptual rendering loss
  - Novel strategy to estimate the non-linear shininess exponent – our proposed ‘mirror image’ representation (perfectly specular version of the object), acts as an absolute reference for the specular shading image, allowing for a more accurate estimation
  - When object normals are available from a depth sensor, the ‘mirror image’ can be unwrapped to also estimate the environment map

3. Pipeline

- Synthetically rendered training dataset with 100,000 samples
- 55 synthetic 3D models augmented with random scaling, orientation and position
- Rendered with uniformly sampled albedo parameters from YUV color space
- 45 indoor environment maps captured in varying lighting conditions - homes, offices, classrooms, auditoriums – randomly rotated before rendering
- Input images augmented with gaussian noise and random background textures

4. Training Data

- Synthetic training dataset with 100,000 samples
- 55 synthetic 3D models augmented with random scaling, orientation and position
- Rendered with uniformly sampled albedo parameters from YUV color space
- 45 indoor environment maps captured in varying lighting conditions - homes, offices, classrooms, auditoriums – randomly rotated before rendering
- Input images augmented with gaussian noise and random background textures

5. Results & Evaluation

- Object segmentation and material estimation results on objects of increasing specularity
- Diffuse-Specular decomposition and Material & Environment Map estimation results

- Qualitative Evaluation: In the ablation study our full method achieves best performance. In comparison to a reflectance-map-based approach, our albedo estimation is slightly worse but the shininess exponent is much more accurate.
- Material retargeting results - We estimate materials from source objects and transfer them to target shapes with consistent lighting by estimating a high-frequency environment map for target shapes from our mirror image layer.
- We also compare against other state-of-the-art methods and obtain more accurate retargeting results.

This work was supported by EPSRC grant CAMERa (EP/M023281/1), ERC Starting Grant CaptReal (335545) and the Max Planck Center for Visual Computing and Communication (MPC-VCC).