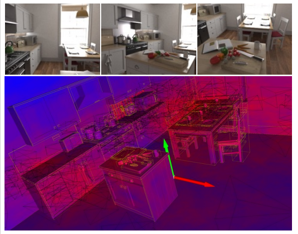
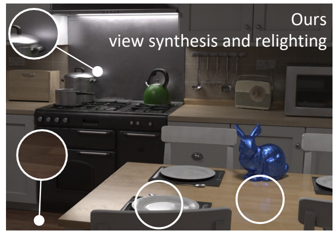
FIPT Factorized Inverse Path Tracing
for Efficient and Accurate Material-Lighting Estimation

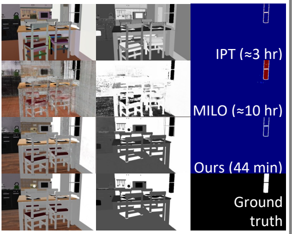
Liwen Wu[1*], Rui Zhu[1*], Mustafa B. Yaldiz[1], Yinhao Zhu[2], Hong Cai[2], Janarbek Matai[2], Fatih Porikli[2], Tzu-Mao Li[1], Manmohan Chandraker[1], Ravi Ramamoorthi[1]
[1]UC San Diego [2]Qualcomm AI Research

Overview


Input views (~200) and geometry



Results and application



Material reflectance Roughness Emission (error map)

Goal:

Given dense HDR observation and geometry of an indoor scene, we seek to efficiently estimate:

- dense materials (BRDF);
- emitters (masks & emission);

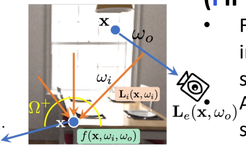
while capturing fine details and specularity.

The rendering equation:

$$\min_{\mathbf{a},\sigma,m,\mathbf{L}_e} \sum_{\mathbf{x},\boldsymbol{\omega}_o} \|\mathbf{L}(\mathbf{x},\boldsymbol{\omega}_o) - \mathbf{L}_{gt}(\mathbf{x},\boldsymbol{\omega}_o)\|_2^2$$

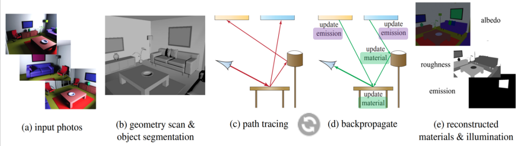
where, $\mathbf{L}(\mathbf{x},\boldsymbol{\omega}_o) = \underbrace{\mathbf{L}_e(\mathbf{x},\boldsymbol{\omega}_o)}_{\text{emission}} + \mathbf{L}_r(\mathbf{x},\boldsymbol{\omega}_o)$

$$\mathbf{L}_r(\mathbf{x},\boldsymbol{\omega}_o) = \int_{\Omega+} \underbrace{\mathbf{L}_i(\mathbf{x},\boldsymbol{\omega}_i)}_{\text{incident radiance}} \underbrace{f(\mathbf{x},\boldsymbol{\omega}_i,\boldsymbol{\omega}_o)}_{\text{BRDF function}} d\boldsymbol{\omega}_i.$$



Inverse Path Tracing (IPT):

- Jointly optimize emission and materials from scratch;
- Differentiable path tracing with gradient descent.



(a) input photos (b) geometry scan & object segmentation (c) path tracing (d) backpropagate (e) reconstructed materials & illumination

- Sensitive to variances.

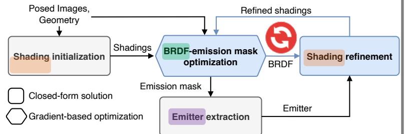

IPT (Piecewise constant materials) MILO (MLP materials) FIPT (MLP materials) Reference

Applications:

- novel view synthesis;
- scene editing (relighting, material editing, object insertion).

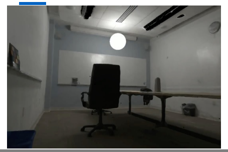
Factorized Inverse Path Tracing (FIPT):

- Factorize and approximate rendering integral into pre-computable terms (i.e. shading, BRDF):
- Analytically solve emitters in the first step, instead of jointly estimating dense emission and BRDF.



emitters are first extracted, then shading and material (BRDF) are updated alternatively.

Resources & demos:



Datasets

- Synthetic indoor datasets;
- Our real-world captures: (1) ~200 frames per-scene; (2) linear response HDR via exposure bracketing.


Conference room Classroom

Method

Factorized light transport:

- Integral-free: (1) optimizable BRDF parameters are moved out of the rendering integrals; (2) then shadings can be pre-computed offline with variance reductions tricks (MIS, denoising, etc.);

$$\mathbf{L}_r(\mathbf{x},\boldsymbol{\omega}_o) = \int_{\Omega+} \underbrace{\mathbf{L}_i(\mathbf{x},\boldsymbol{\omega}_i)}_{\text{incident radiance}} \underbrace{f(\mathbf{x},\boldsymbol{\omega}_i,\boldsymbol{\omega}_o)}_{\text{BRDF function}} d\boldsymbol{\omega}_i.$$
$$= \mathbf{k}_d \mathbf{L}_d(\mathbf{x})$$
$$+ \mathbf{k}_s \mathbf{L}_s^0(\mathbf{x},\boldsymbol{\omega}_o,\sigma) + \mathbf{L}_s^1(\mathbf{x},\boldsymbol{\omega}_o,\sigma)$$

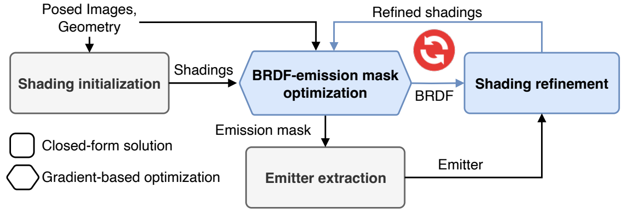
where, $\mathbf{L}_d(\mathbf{x}) = \int_{\Omega+} \mathbf{L}_i(\mathbf{x},\boldsymbol{\omega}_i) \frac{(\mathbf{n}\cdot\boldsymbol{\omega}_i)_+}{\pi} d\boldsymbol{\omega}_i$ ◀ diffuse shading

$$\mathbf{L}_s^0(\mathbf{x},\boldsymbol{\omega}_o,\sigma) = \int_{\Omega+} \mathbf{L}_i(\mathbf{x},\boldsymbol{\omega}_i) \frac{F_0 DG}{4(\mathbf{n}\cdot\boldsymbol{\omega}_o)} d\boldsymbol{\omega}_i$$
$$\mathbf{L}_s^1(\mathbf{x},\boldsymbol{\omega}_o,\sigma) = \int_{\Omega+} \mathbf{L}_i(\mathbf{x},\boldsymbol{\omega}_i) \frac{F_1 DG}{4(\mathbf{n}\cdot\boldsymbol{\omega}_o)} d\boldsymbol{\omega}_i,$$

◀ specular shadings

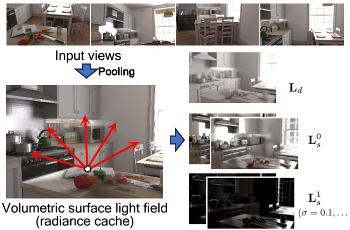
- Pre-compute specular shadings with 6 pre-defined roughness levels σ_k:
  - Given estimated roughness, shading can be linearly interpolated from pre-baked shadings.
  $\mathbf{L}_s^*(\cdot,\sigma) \approx \text{lerp}(\{\mathbf{L}_s^*(\cdot,\sigma_k)|\sigma_k \in \text{linspace(0,1,6)}\},\sigma)$
- Alternate optimization based on factorization:
  Initialize/re-bake $\{\mathbf{L}_d, \mathbf{L}_s^0, \mathbf{L}_s^1\}$ Optimize $\{\mathbf{L}_e, \mathbf{k}_d, \mathbf{k}_s, \sigma\}$

Pipeline:



- **Shading initialization**

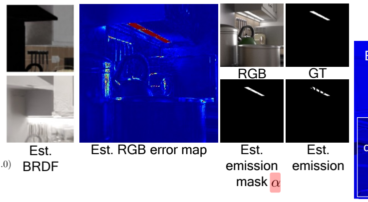
To initialize the shadings, we pool the input views onto a volumetric surface light field. Querying it as an approximation of the global incident illumination allows us to estimate the initial diffuse and specular shading integrals for each input view.



Volumetric surface light field (radiance cache)

- **BRDF and emission estimation**

Given the initial shadings we solve for BRDF and emission. The materials are encoded in a coordinate based MLP then updated by gradient descent. We optimize without emission term but an emission mask with sparsity constraint. This compensates error near emissive surface and tell which part is an emitter.



[IPT] Inverse Path Tracing for Joint Material and Lighting Estimation'19
[Li22] Physically-Based Editing of Indoor Scene Lighting from a Single Image'22
[MILO] MILO: Multi-bounce Inverse Rendering for Indoor Scene with Light-emitting Objects'23
[FVP] Free-viewpoint Indoor Neural Relighting from Multi-view Stereo'21
References

- Other novelty:

Neural representation of materials:

$$\mathbf{k}_d, \mathbf{k}_s, \sigma = \mathbf{MLP}_{\Theta_1}(\mathbf{hashgrid}(\mathbf{x}))$$
$$\alpha = \mathbf{MLP}_{\Theta_2}(\gamma(\mathbf{x}))$$

Optimize material and emission with sparsity priors:

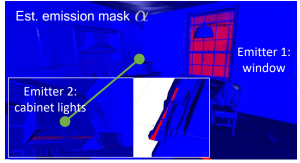
$$\underset{\Theta_1,\Theta_2}{\text{argmin}} |(1-\alpha)(\mathbf{k}_d\mathbf{L}_d + \mathbf{k}_s\mathbf{L}_s^0 + \mathbf{L}_s^1 - \mathbf{L}_{gt})|_2^2 + \underbrace{\lambda|\alpha|^2}_{\text{emission mask prior}} + \underbrace{\lambda' l_\sigma}_{\text{roughness regularizer}}$$

- Leads to more piece-wise constant material estimation;
- Enables fast error-driven emitter estimation.

- **Error-driven emitter estimation**

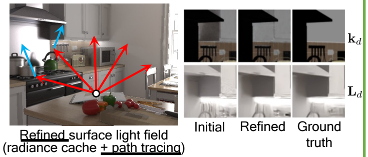
Optimization without emission terms produces distinctive error near emissive surfaces (1st column). By jointly optimizing an emission mask to cancel this error, the emitter can be found by checking the mask's response (works even for tiny emitters), and its emission value can be obtained by median-pooling the RGBs from training pixels.

$$\mathbf{L}_e = \begin{cases} \text{median}_{\mathbf{x}}\ \mathbf{L}_{gt}(\mathbf{x}) & \alpha > 0.01 \\ 0 & \text{otherwise} \end{cases}$$



- **Shading refinement**

The wall cabinet's diffuse reflectance estimation is initially darker than ground truth, owing to the excessive incident light received from the range hood that reflects non-diffuse light (2nd column). The artifacts are reduced by growing the path (blue) for the specular surface according to the optimized BRDF (1st column), which gives more accurate shadings that can be used to further refine the BRDF (3rd column).



Refined surface light field (radiance cache + path tracing)

Compare (initial VS refined global illumination)

Evaluation: synthetic scenes

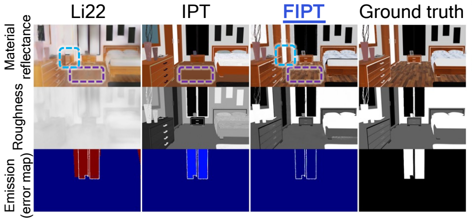
	Method	Bathroom	Bedroom	Livingroom	Kitchen
View synthesis	FVP	23.38	20.49	24.63	20.77
	IPT	14.76	21.85	23.87	19.94
	MILO	20.62	20.25	24.47	18.09
	FIPT	25.42	29.84	**30.86**	**25.38**
	FIPT-sem	**25.76**	**29.89**	30.84	25.27
Relight	Li22	22.86	23.20	19.83	21.76
	FVP	23.72	24.11	19.51	23.31
	IPT	20.61	28.16	27.26	27.28
	MILO	14.97	23.39	22.10	19.62
	FIPT	**31.28**	36.64	**31.56**	**29.13**
	FIPT-sem	31.03	**36.69**	30.82	28.79

PSNR of view synthesis and relighting

	Method	$\mathbf{k}_d$	$\mathbf{a}'$ PSNR↑	σ	$\mathbf{L}_e$ IoU↑	$\mathbf{L}_e$ logL2↓
Bathroom	Li22	19.92	15.78	13.77	0.45	1.35
	NeILF	10.12	9.01	14.82	-	-
	IPT	22.43	18.59	14.69	0.33	1.09e-1
	MILO	11.83	9.80	5.56	0.05	5.60e-1
	FIPT	**30.13**	**25.28**	**28.79**	**0.63**	**3.18e-2**
	FIPT-sem	27.81	24.00	21.84	**0.63**	**3.18e-2**
Bedroom	Li22	21.87	17.18	12.12	0.34	2.78
	NeILF	14.88	12.42	11.30	-	-
	IPT	29.39	22.46	13.33	0.92	4.01e-3
	MILO	23.65	15.16	15.42	0.08	1.59e-2
	FIPT	**31.10**	**29.41**	23.19	**0.96**	4.95e-4
	FIPT-sem	31.00	28.45	**25.23**	**0.96**	**4.93e-4**
Livingroom	Li22	17.25	15.32	12.72	0.17	3.61
	NeILF	12.34	10.97	13.45	-	-
	IPT	21.24	19.01	11.77	0.90	6.08e-3
	MILO	22.88	18.39	13.98	0.06	1.39e-2
	FIPT	28.86	**28.70**	**32.48**	**0.95**	**8.06e-4**
	FIPT-sem	**29.09**	28.62	25.15	**0.95**	8.09e-4
Kitchen	Li22	18.14	14.54	10.82	0.43	1.41
	NeILF	12.63	9.96	10.64	-	-
	IPT	25.68	21.61	11.84	0.83	1.08e-2
	MILO	18.25	13.86	12.56	0.10	8.28e-2
	FIPT	33.07	**27.53**	**29.24**	**0.91**	**1.54e-3**
	FIPT-sem	**33.25**	27.38	21.70	**0.91**	**1.54e-3**

BRDF-emission comparison

FIPT-sem: FIPT + semantic segmentation to encourage piece-wise smoothness of BRDF



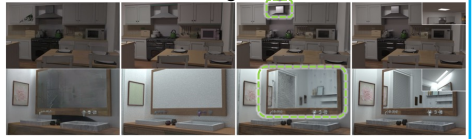
BRDF-emission estimation

⇙ Our method successfully reconstructs BRDF (1st and 2nd rows) and emission (3rd row) with high frequency details and less ambiguity (e.g. floor texture and mirror). Emission estimation is shown as error heatmaps.

⤋ We better reproduce specularity in lighting and materials (e.g. the range hood and mirror).

View synthesis

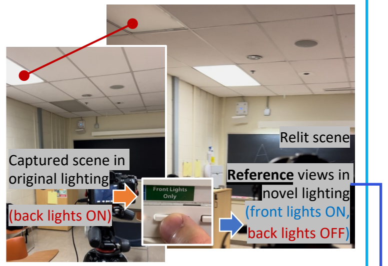
Relighting



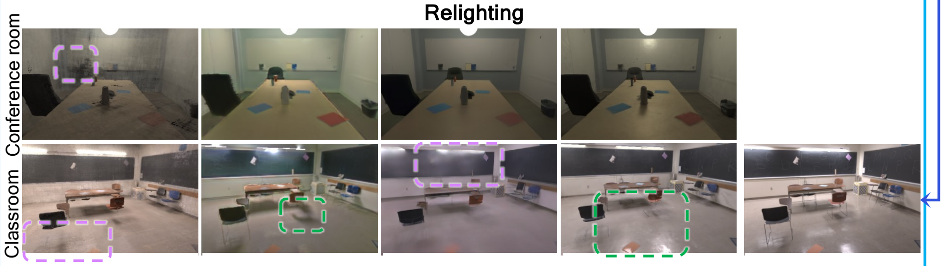
Evaluation: real scenes

⇛ We test all methods in the task of light editing, by inputting images of original lighting, and rendering images under novel lighting by inserting virtual front lights, in novel views. We can then use our captured images of novel lighting as reference (or pseudo ground truth).

⤋ As we may observe here, our method is the closest to the reference, with shadows under the desk and specular highlights on the floor decently reproduced, with minimal artifacts.



Relighting



MILO Li22 FVP FIPT Reference