Learning to Predict Indoor Illumination from a Single Image

Marc-André Gardner, Kalyan Sunkavalli, Ersin Yumer, Xiaohui Shen,
Emiliano Gambaretto, Christian Gagné, Jean-François Lalonde

Université Laval, Québec, Canada
Adobe Systems Inc., San Jose, USA

Illumination is key

Illumination is key

Current approaches and limitations

Calibration objects

[Debevec, 1998]

Current approaches and limitations

Calibration objects

[Debevec, 1998]

Specialized hardware

[Tocci, 2011], [Manakov, 2013]

Current approaches and limitations

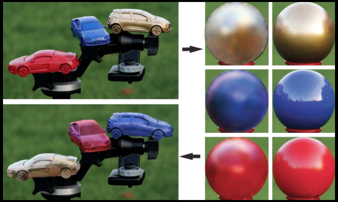
Calibration objects

[Debevec, 1998]

Specialized hardware

[Tocci, 2011], [Manakov, 2013]

Scene knowledge



[Rematas, 2015]

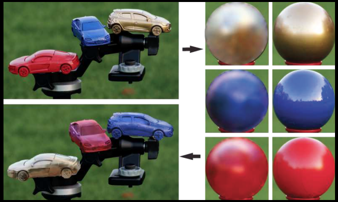
Our approach

No calibration

Specialized hardware

[Tocci, 2011], [Manakov, 2013]

Scene knowledge



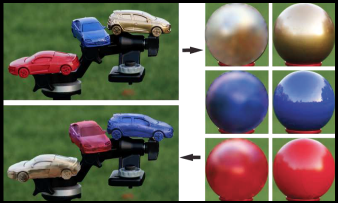
[Rematas, 2015]

Our approach

No calibration

Any camera

Scene knowledge



[Rematas, 2015]

Our approach

No calibration

Any camera

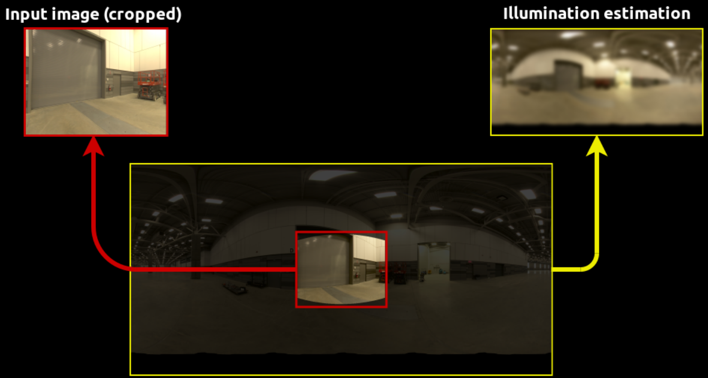
No prior knowledge

Our goal

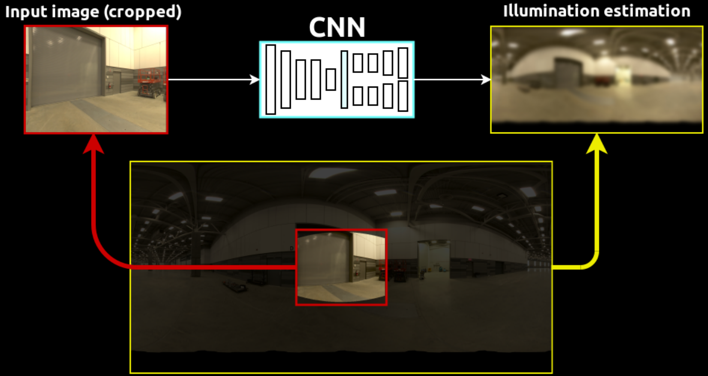
Given a single indoor LDR image, recover a whole HDR environment map

End-to-end learning approach

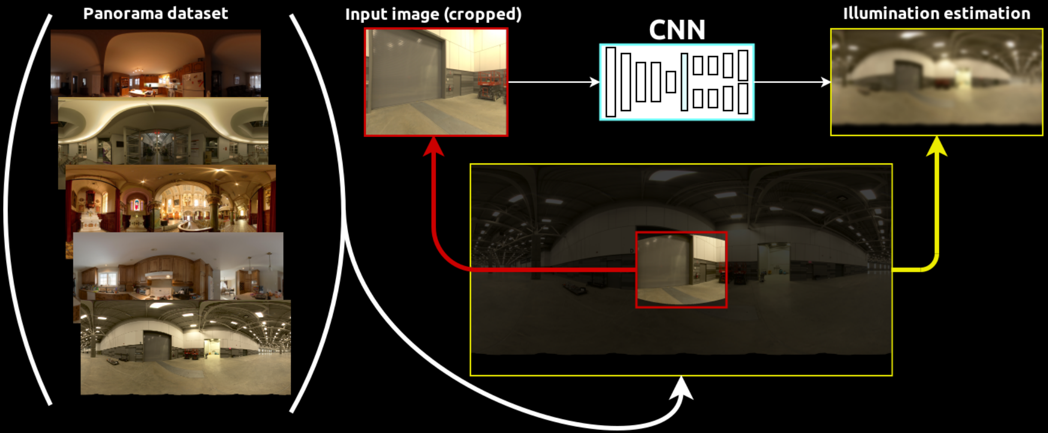
End-to-end learning approach



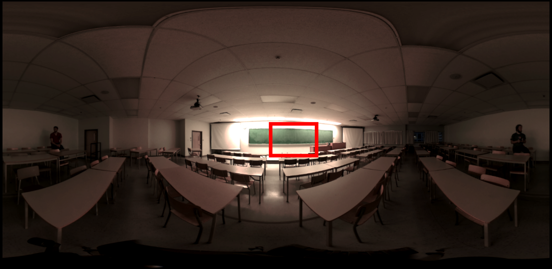
End-to-end learning approach



End-to-end learning approach

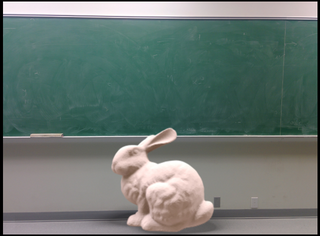


Spatially-varying lighting



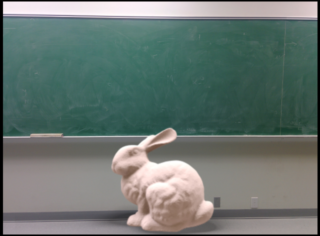
Spatially-varying lighting

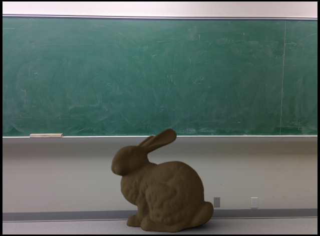
Actual lighting



Spatially-varying lighting

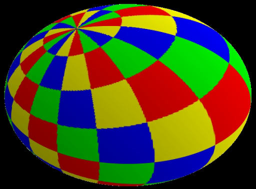
Actual lighting $\neq$ Panorama lighting



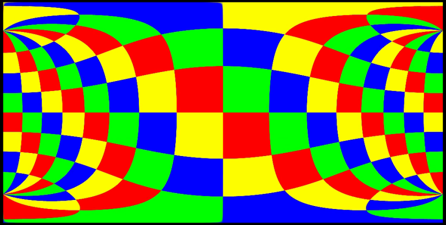


Spatially-varying lighting: the *warp* operator

Sphere



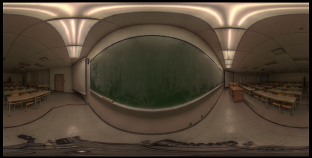
Latlong projection



Panorama example

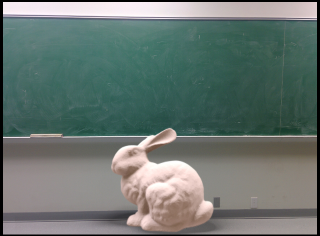
Spatially-varying lighting: warping in practice

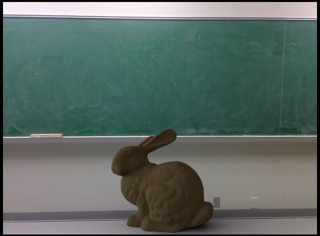
Spatially-varying lighting: warping in practice



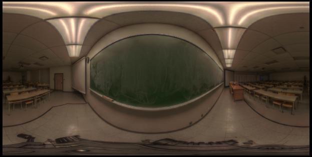
Actual lighting

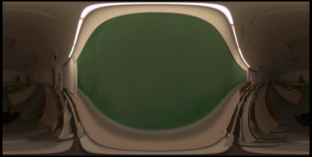
Panorama lighting





Spatially-varying lighting: warping in practice

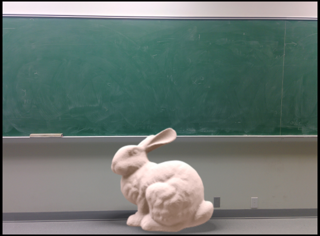


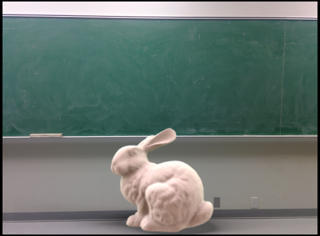


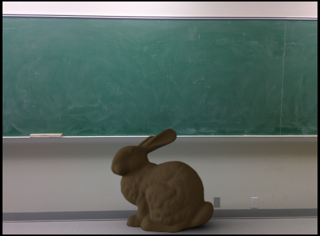
Actual lighting

Estimated lighting

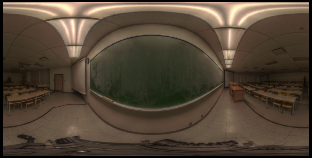
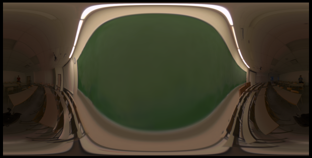
Panorama lighting



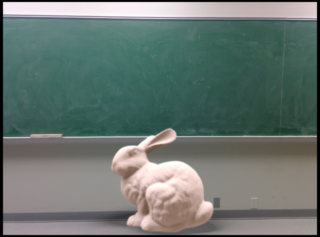
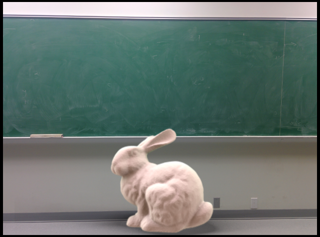
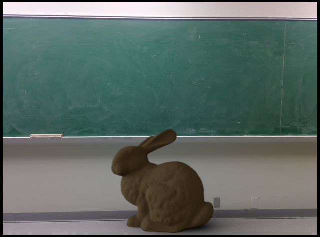




Spatially-varying lighting: warping in practice


Actual lighting | Estimated lighting | Panorama lighting

- Automatic approach (no user intervention needed)

Our goal

Given a single indoor LDR image, recover a whole HDR environment map

☑ Spatially-varying lighting

Recovering light positions: dataset creation

- SUN360 dataset

Recovering light positions: dataset creation

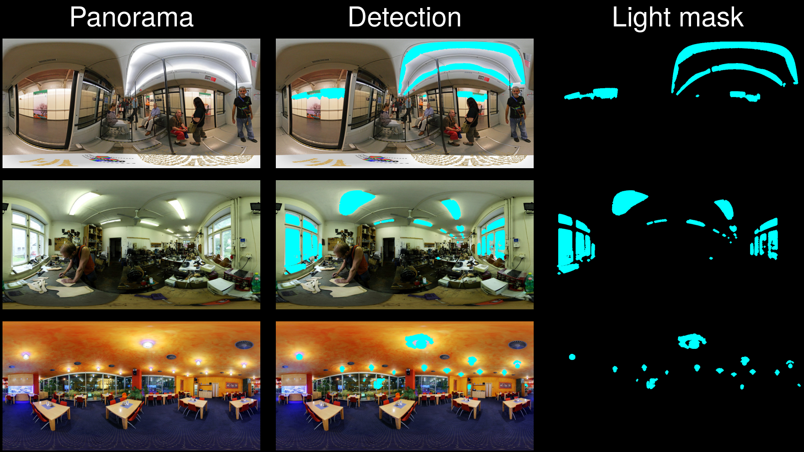
- SUN360 dataset
- No HDR!

Recovering light positions: dataset creation

- SUN360 dataset
- No HDR!
- HOG+SVM+CRF
  light mask detector

Recovering light positions: dataset creation

- SUN360 dataset
- No HDR!
- HOG+SVM+CRF *light mask* detector



Our goal

Given a single indoor LDR image, recover a whole HDR environment map

- ☑ Spatially-varying lighting
- ☑ Light mask creation

Recovering light positions: network architecture

- End-to-end deep neural network

Recovering light positions: network architecture

- End-to-end deep neural network
  - Takes a picture as input

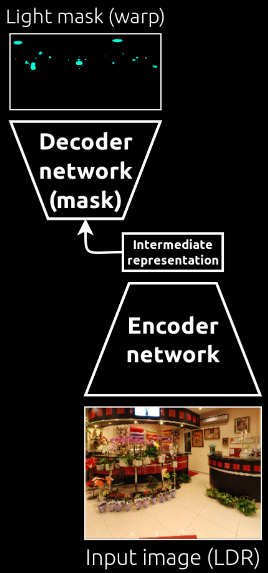
Input image (LDR)

Recovering light positions: network architecture

- End-to-end deep neural network
  - Takes a picture as input
  - *Encodes* it

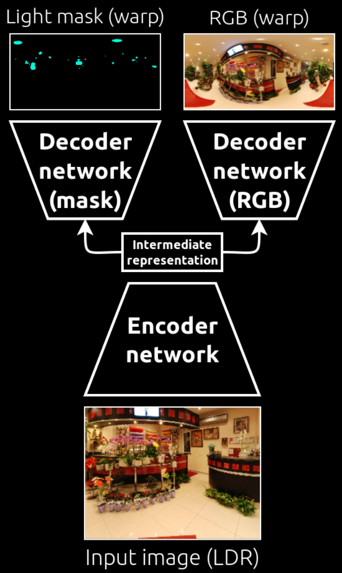
Recovering light positions: network architecture

- End-to-end deep neural network
  - Takes a picture as input
  - *Encodes* it
  - Decodes to get:
    1. Light mask



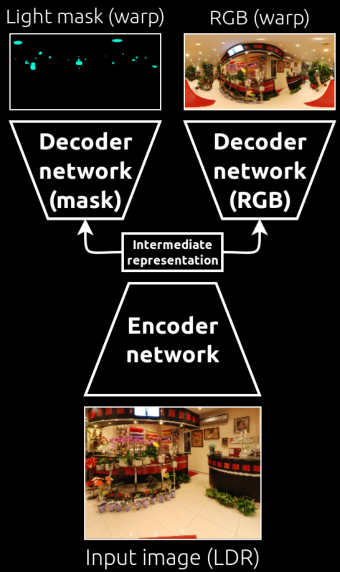
Recovering light positions: network architecture

- End-to-end deep neural network
  - Takes a picture as input
  - *Encodes* it
  - Decodes to get:
    1. Light mask
    2. RGB panorama



Recovering light positions: network architecture

- End-to-end deep neural network
  - Takes a picture as input
  - *Encodes* it
  - Decodes to get:
    1. Light mask
    2. RGB panorama
- In more details:
  - Encoder: ResNet blocks
  - Decoders: deconv layers
  - No pooling (striding)
  - Adam optimizer
  - 38M parameters

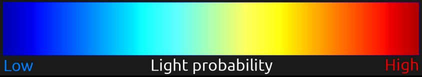


Our goal

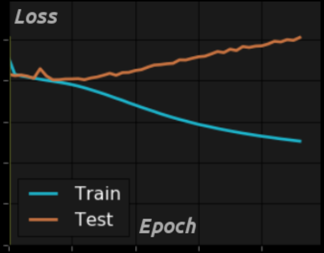
Given a single indoor LDR image, recover a whole HDR environment map

- ☑ Spatially-varying lighting
- ☑ Light mask creation
- ☑ Network architecture

Training



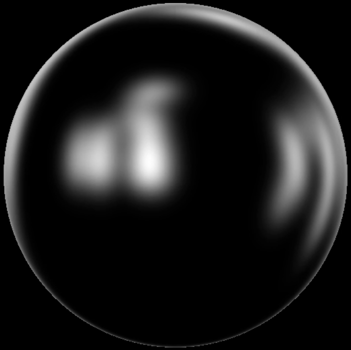
- L2 loss on the RGB, cross-entropy on the light mask



Training (loss function)

- We introduce a tunable render loss function
- L2 loss on the *filtered* masks

$$\mathcal{F}(\mathbf{p}, i, e) = \frac{1}{K_i} \sum_{\omega \in \Omega_i} \mathbf{p}(\omega) s(\omega) \left(\omega \cdot n_i\right)^{\alpha e}$$



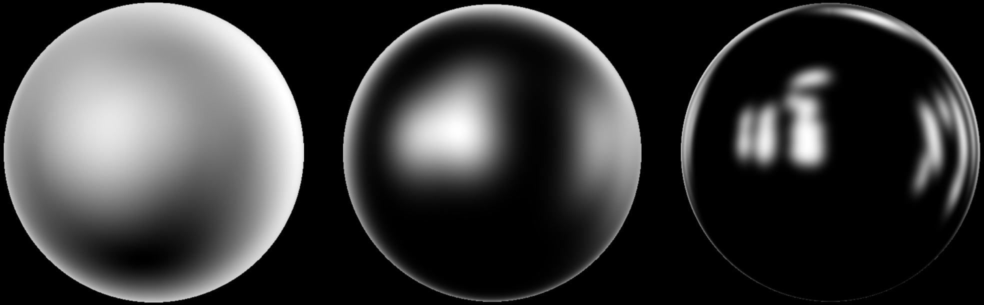
$\alpha e = 50$

Training (loss function)

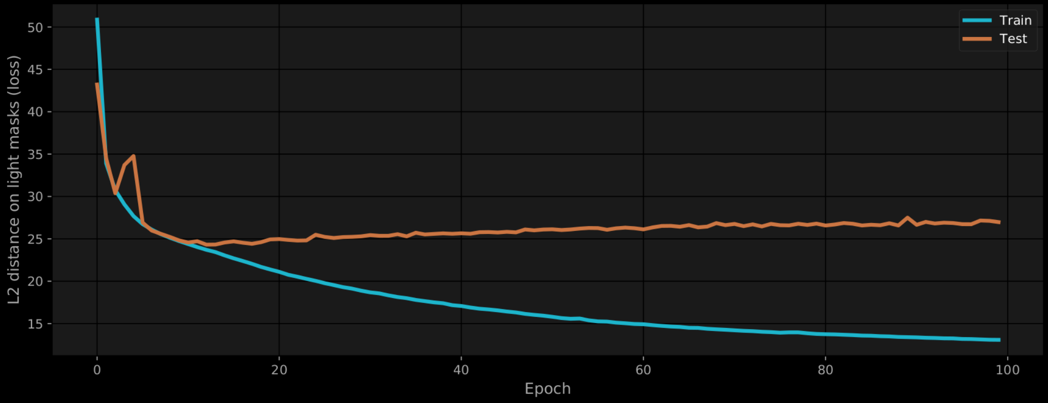
- We introduce a tunable render loss function
- L2 loss on the *filtered* masks

$$\mathcal{F}(\mathbf{p}, i, e) = \frac{1}{K_i} \sum_{\omega \in \Omega_i} \mathbf{p}(\omega)\mathbf{s}(\omega)\,(\omega \cdot n_i)^{\alpha e}$$

$\alpha e = 1$ (diffuse) $\alpha e = 10$ $\alpha e = 80$ (glossy)



Training (loss function)



Our goal

Given a single indoor LDR image, recover a whole HDR environment map

- ☑ Spatially-varying lighting
- ☑ Light mask creation
- ☑ Network architecture
- ☑ Loss function

Results (light position)

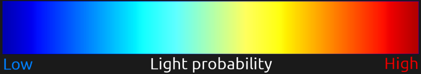
Low Light probability High

Results (light position)

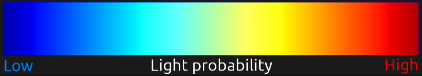
Results (light position)

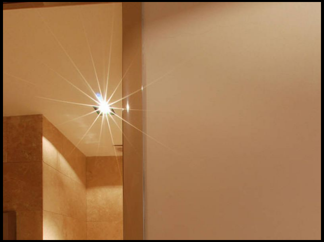
Low Light probability High

Results (light position)

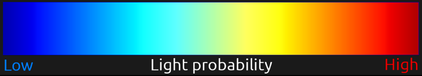


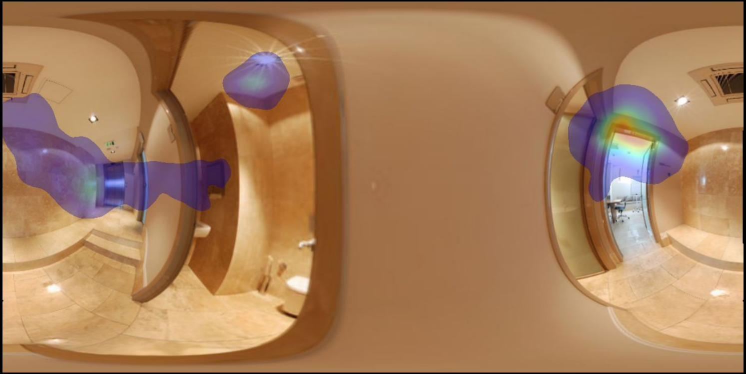
Results (light position)





Results (light position)



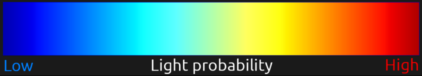


Results (light position)

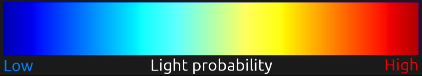
Low Light probability High

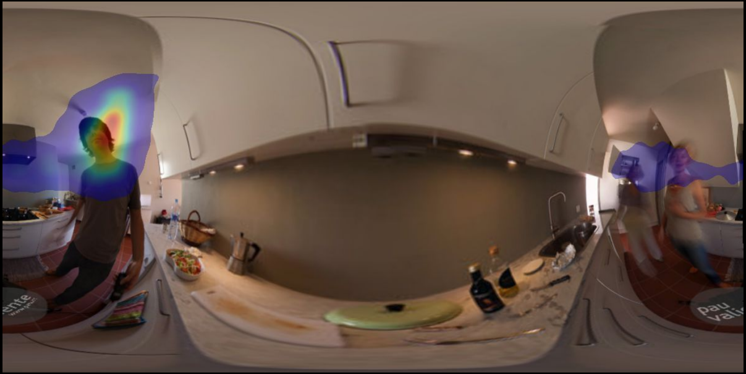
Results (light position)

Results (light position)



Results (light position)



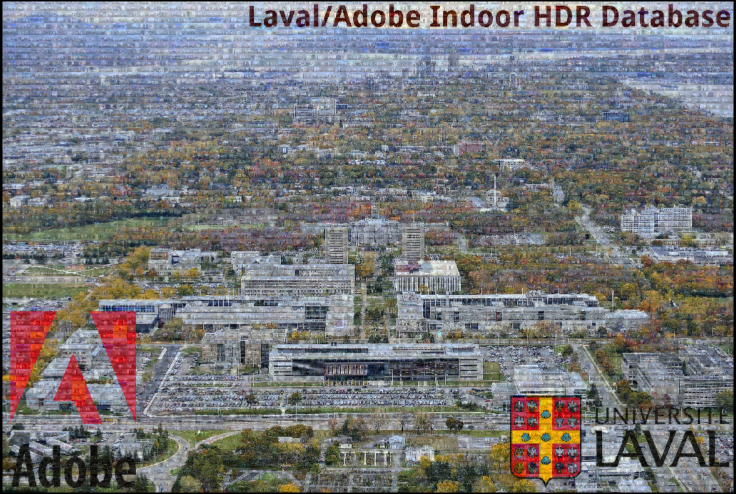


Given a single indoor LDR image, recover a whole HDR environment map

- ✓ Spatially-varying lighting
- ✓ Light mask creation
- ✓ Network architecture
- ✓ Loss function
- ✓ Light position

Adding HDR data

- New HDR panorama dataset: `indoor.hdrdb.com`
  - 2100 indoor scenes, fully HDR, high-res

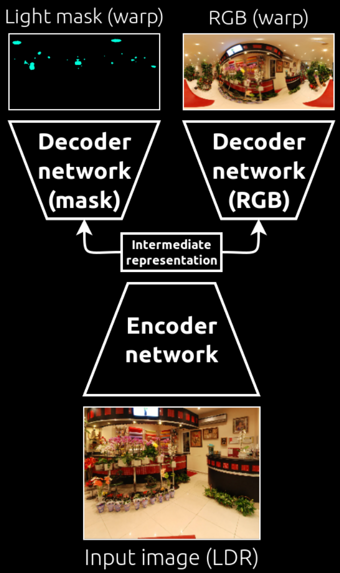


Categories:

- bedroom
- living room
- machine shop
- church
- elevator
- bathroom
- lobby
- laboratory
- auditorium
- hallway
- classroom
- kitchen
- basement
- kids room
- museum
- grocery store
- factory
- storage room
- sports facility
- tunnel
- staircase
- office lobby
- conference room
- etc...

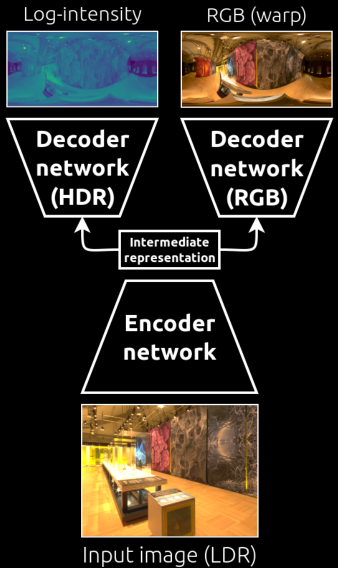
Towards light intensity recovery

- *Fine-tuning* on HDR data



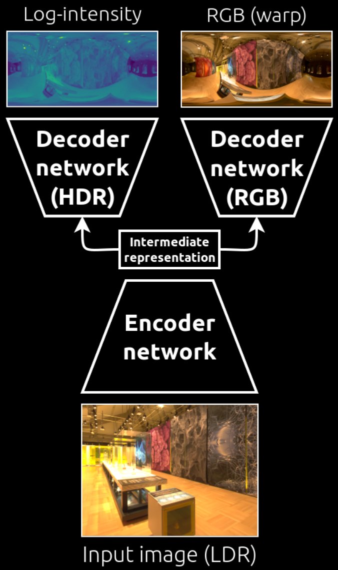
Towards light intensity recovery

- *Fine-tuning* on HDR data
  - RGB head → idem (ambient term)
  - Light mask → *log-intensity*



Towards light intensity recovery

- *Fine-tuning* on HDR data
  - RGB head → idem (ambient term)
  - Light mask → *log-intensity*
- Render loss used for log-intensity head



Results (light intensity and RGB)

Results (light intensity and RGB)

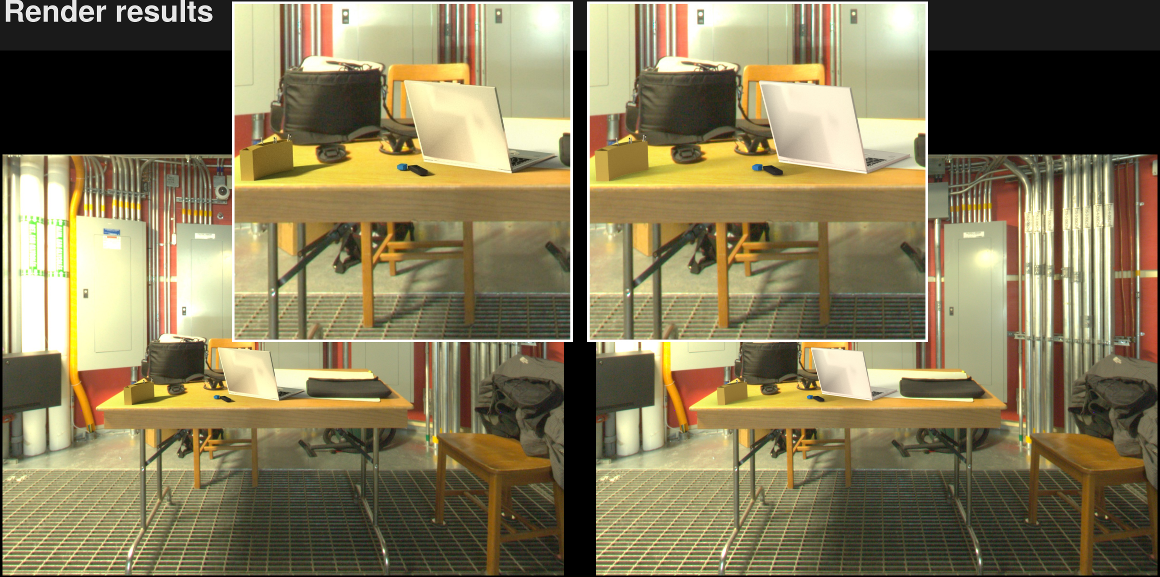
Results (light intensity and RGB)

Our goal

Given a single indoor LDR image, recover a whole HDR environment map

- ☑ Spatially-varying lighting
- ☑ Light mask creation
- ☑ Network architecture
- ☑ Loss function
- ☑ Light position
- ☑ HDR fine-tuning

Render results

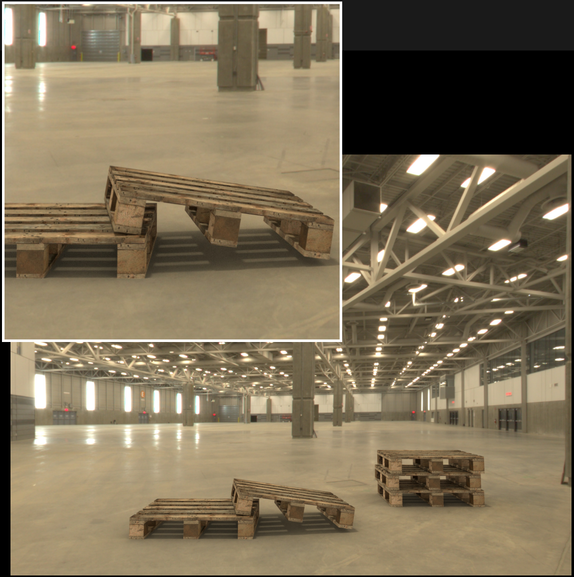


Ground truth

Prediction

Render results

Ground truth



Prediction

Render results

Ground truth

Prediction

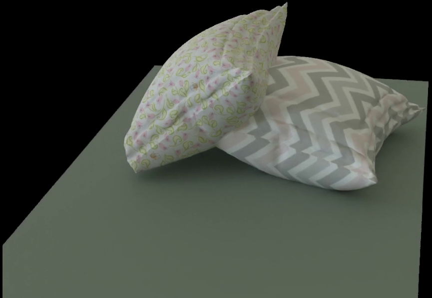
Render results

Ground truth

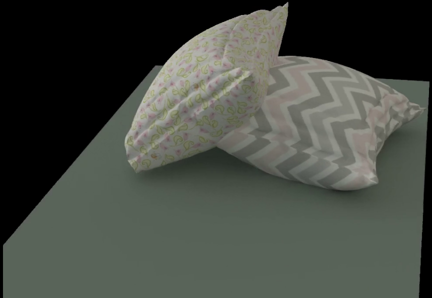
Prediction

Results (render)

Ground truth



Prediction



User study

- A/B comparison against ground truth lighting

User study

- A/B comparison against ground truth lighting
- 105 users, 20 scenes

Pair #3

Which image from Pair #3 looks MORE realistic? *

- ◯ The left one
- ◯ The right one
- ◯ They both look equally plausible

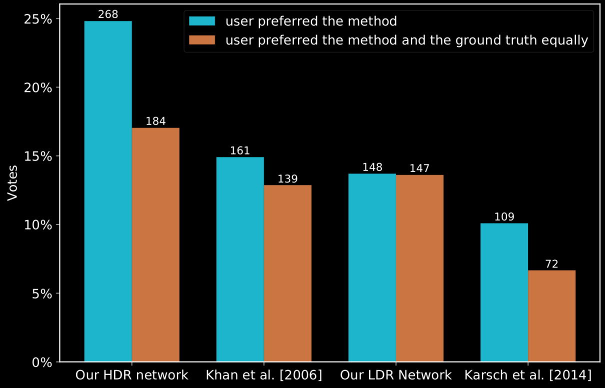
Pair #4

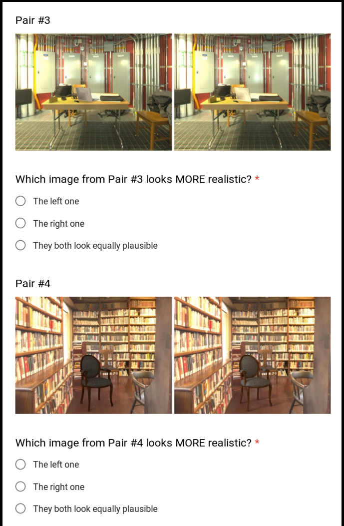
Which image from Pair #4 looks MORE realistic? *

- ◯ The left one
- ◯ The right one
- ◯ They both look equally plausible

User study

- A/B comparison against ground truth lighting
- 105 users, 20 scenes
- Our approach outperforms state-of-the-art methods





Conclusion

Main contribution: framing lighting estimation as end-to-end learning

- New HDR panorama dataset

Conclusion

Main contribution: framing lighting estimation as end-to-end learning

- New HDR panorama dataset
- Novel warping operator and render loss function

Conclusion

Main contribution: framing lighting estimation as end-to-end learning

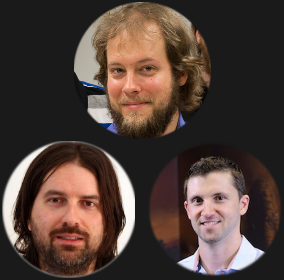
- New HDR panorama dataset
- Novel warping operator and render loss function

Conclusion

Main contribution: framing lighting estimation as end-to-end learning

- New HDR panorama dataset
- Novel warping operator and render loss function

Thank you!



UNIVERSITÉ LAVAL

Adobe

More info: `jflalonde.ca/projects/deepIndoorLight`

HDR dataset: `indoor.hdrdb.com`