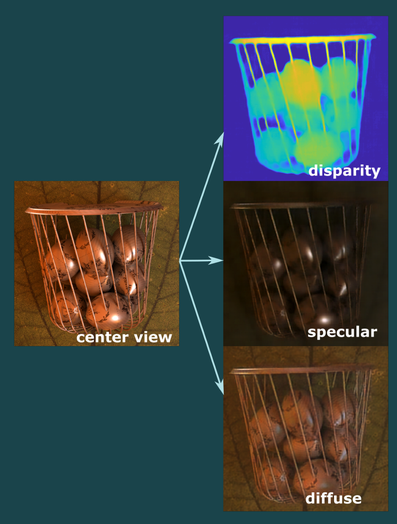
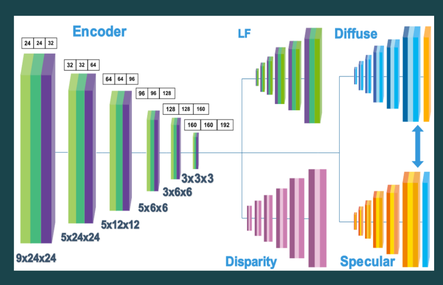
# Light Field Intrinsics With a Deep Encoder-Decoder Network

**Anna Alperovich, Ole Johannsen, Michael Strecke and Bastian Goldluecke**
University of Konstanz
https://www.cvia.uni-konstanz.de/

## Contributions

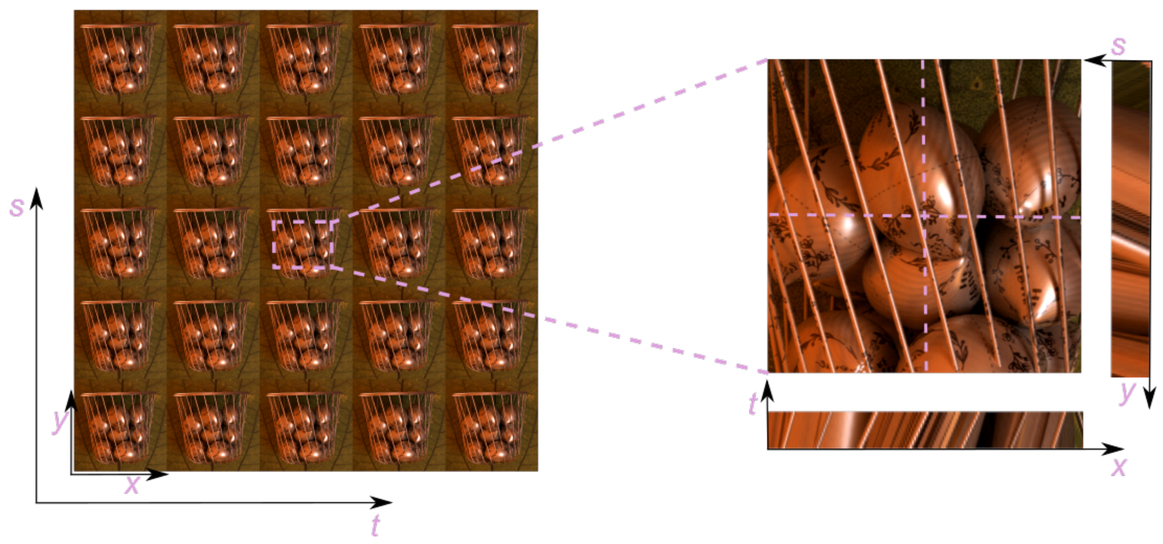
- Encoder-decoder network for disparity regression and reflection separation in light fields
- Supervised and unsupervised training
- 3D convolutions to compute features over the vertical and horizontal stack
- Synthetic light field data with ground truth intrinsic components and disparity




## Light field structure

Defined on 4D ray space $\mathcal{R} = \Pi \times \Omega$, which parametrizes rays $\boldsymbol{r} = (x, y, s, t)$ by their intersection coordinates with two planes $\Pi$ and $\Omega$, [5]. Intersection with the focal plane $\Pi$ gives view point coordinates $(s, t)$, while the image plane $\Omega$ denotes image coordinates $(x, y)$.



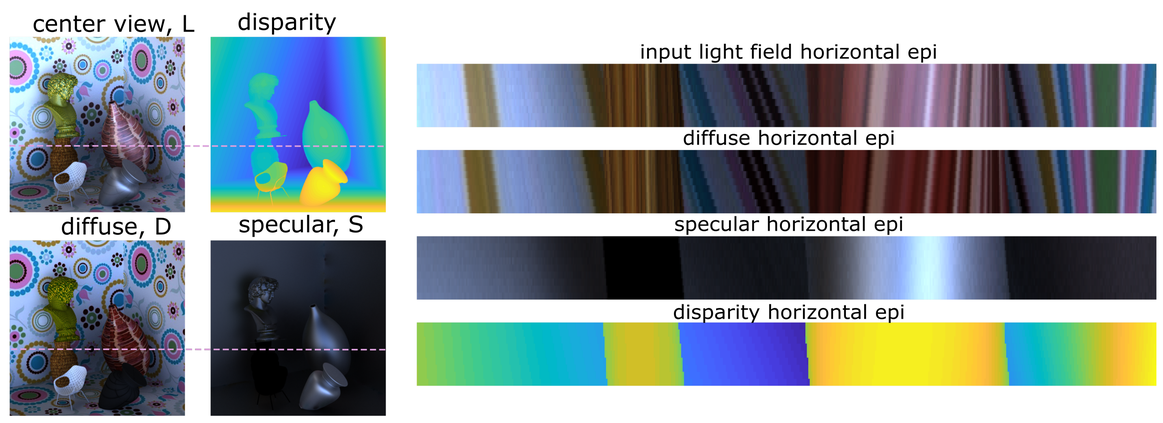

## Dichromatic model

$$L = D + S$$

The input light field $L$ can be additively separated into its diffuse and specular components $D$ and $S$, [6], [10].

**Diffuse component** and the disparity correspond to the same projections of the same 3D points and share the same pattern.

**Specular component** follows the specular flow [9], which depends on the local surface geometry and view point change.

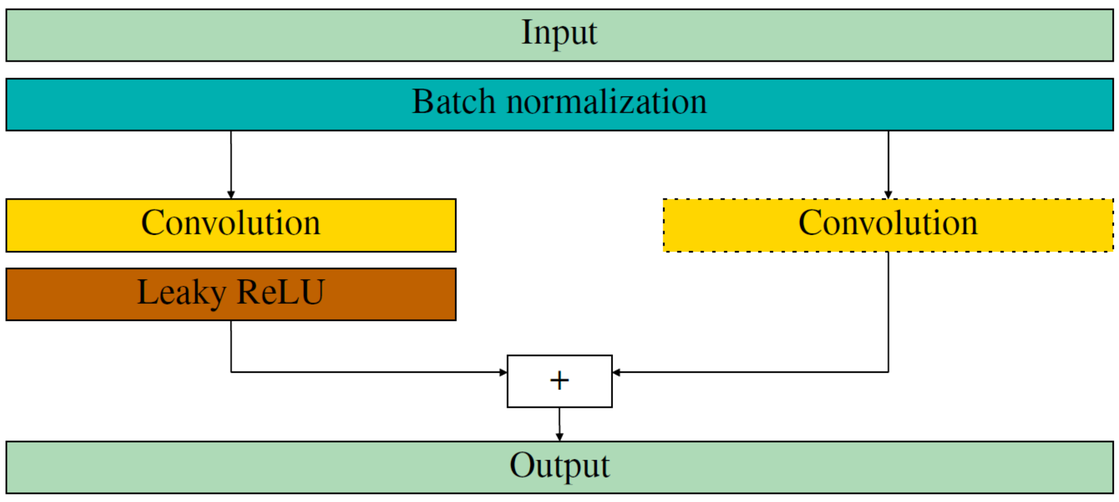



## Data


Generated based on Blender plugin provided by [2]

- 171 scenes
- 321 textures
- 109 environmental maps[3]
- 36 pre-built scenes with objects from Chocofur[1] and The British Museum[2]
- Randomized position, amount of specularity, texture and color

## Residual block

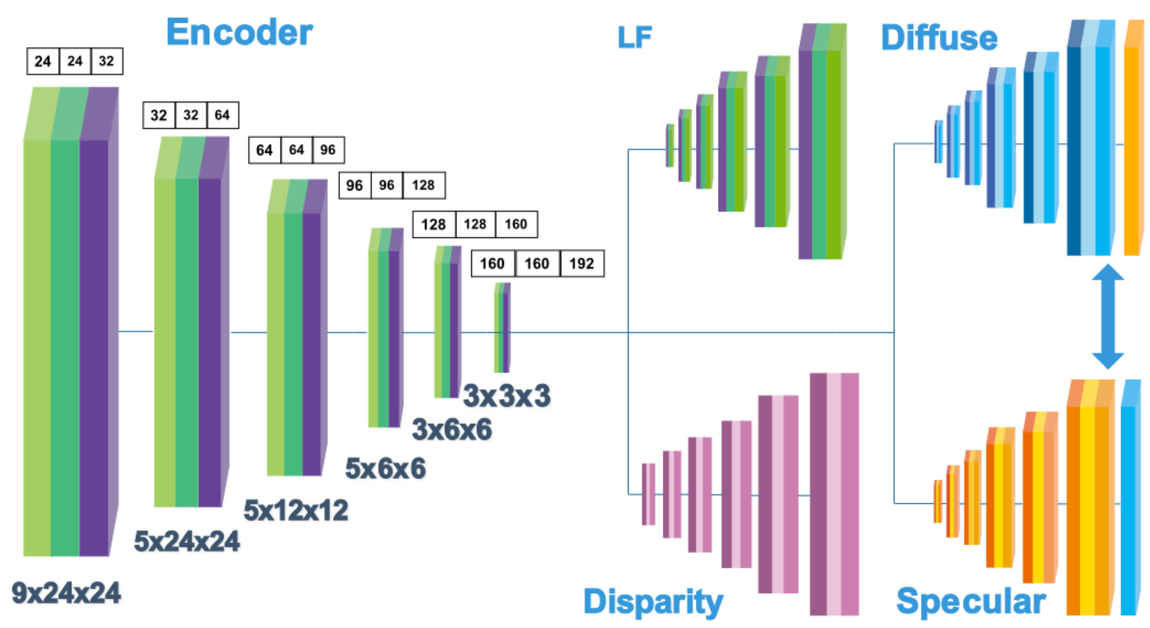



After **batch normalization**, a first path leads through a (possibly strided) **convolution** layer and a **leaky ReLU**.

A second path either keeps the input, or passes it through a strided **convolution** in case it needs to be resampled.

Both paths are added together to produce the final output.

## Network architecture



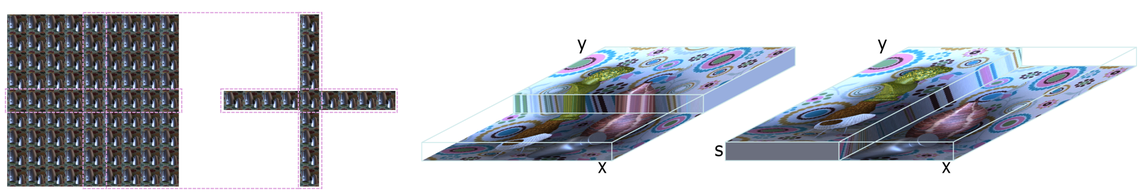

- Paths:
  - One encoding
  - Four decoding
  - Six groups of three residual blocks each
  - Disparity is a 2D decoder
  - Reconstructed light field, diffuse and specular decoders are 3D
- Block types:
  - depth and resolution the same
  - reduced resolution, while increased feature depth
- Network operates on the horizontal and vertical EPI stacks in parallel
- The feature output is joined on the bottom layer
- Compression rate is $8.3\%$

This work was supported by the ERC Starting Grant
"Light Field Imaging and Analysis" (LIA 336978, FP7-2014).
Presented at the International Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, June 2018.

# Light Field Intrinsics With a Deep Encoder-Decoder Network

**Anna Alperovich, Ole Johannsen, Michael Strecke and Bastian Goldluecke**
University of Konstanz
https://www.cvia.uni-konstanz.de/

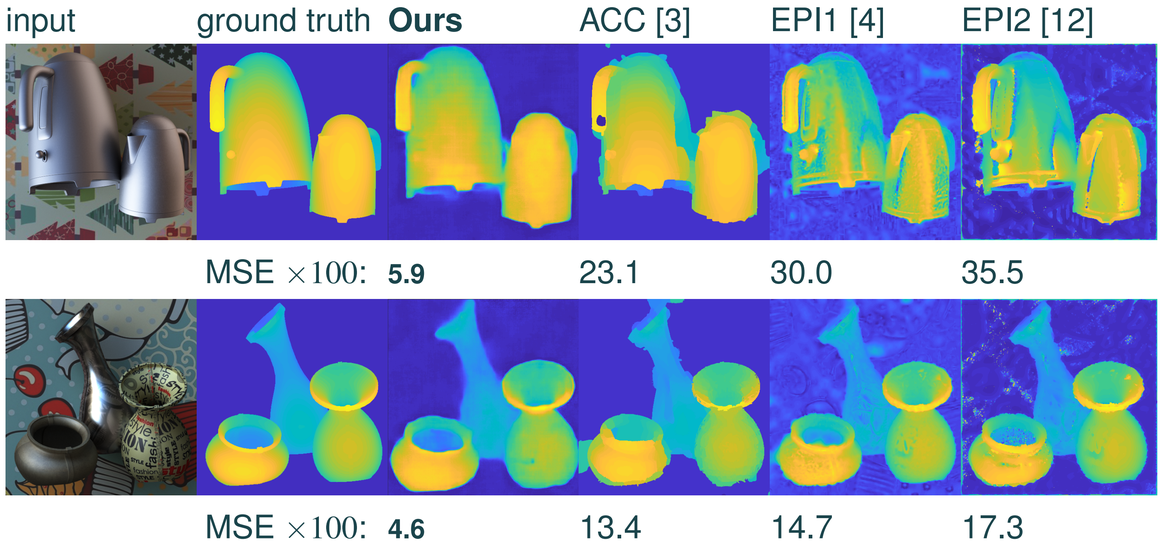
## Network inputs




- crosshair-shaped subset of 17 views
- patch size $9 \times 48 \times 48$
- $\approx 160,000$ patches

## Comparisons on the synthetic light fields

### disparity



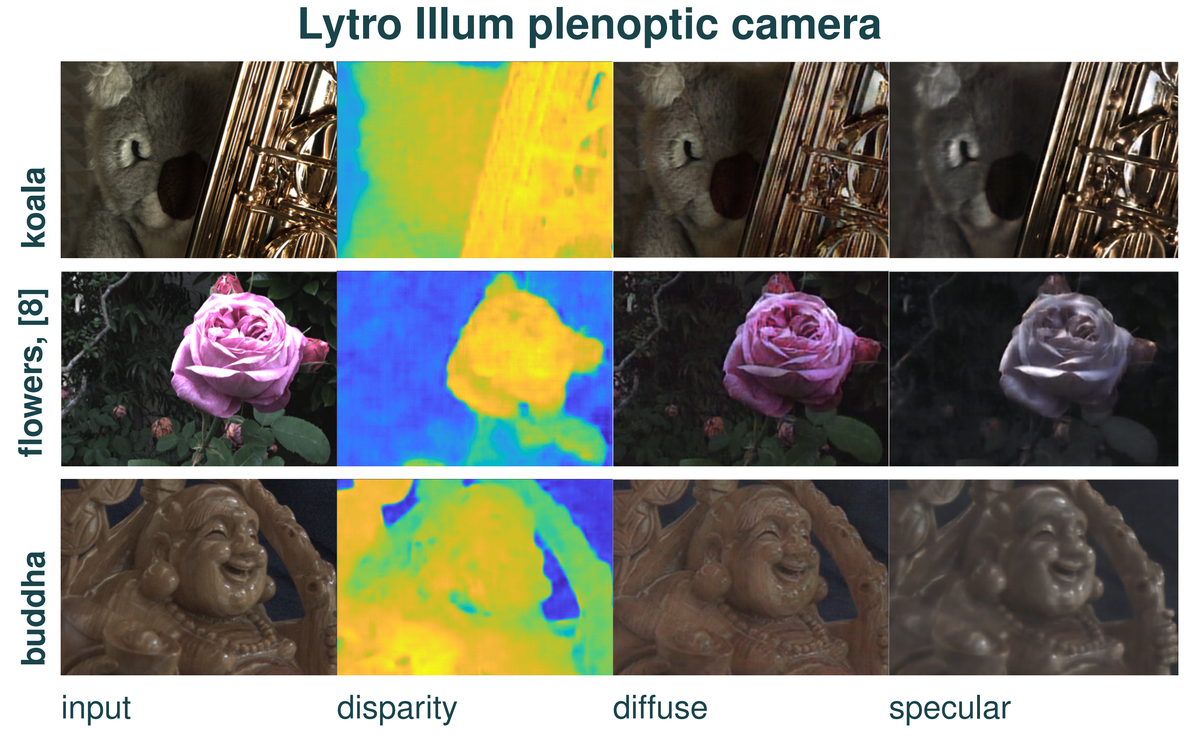

### reflection separation


### numerical results

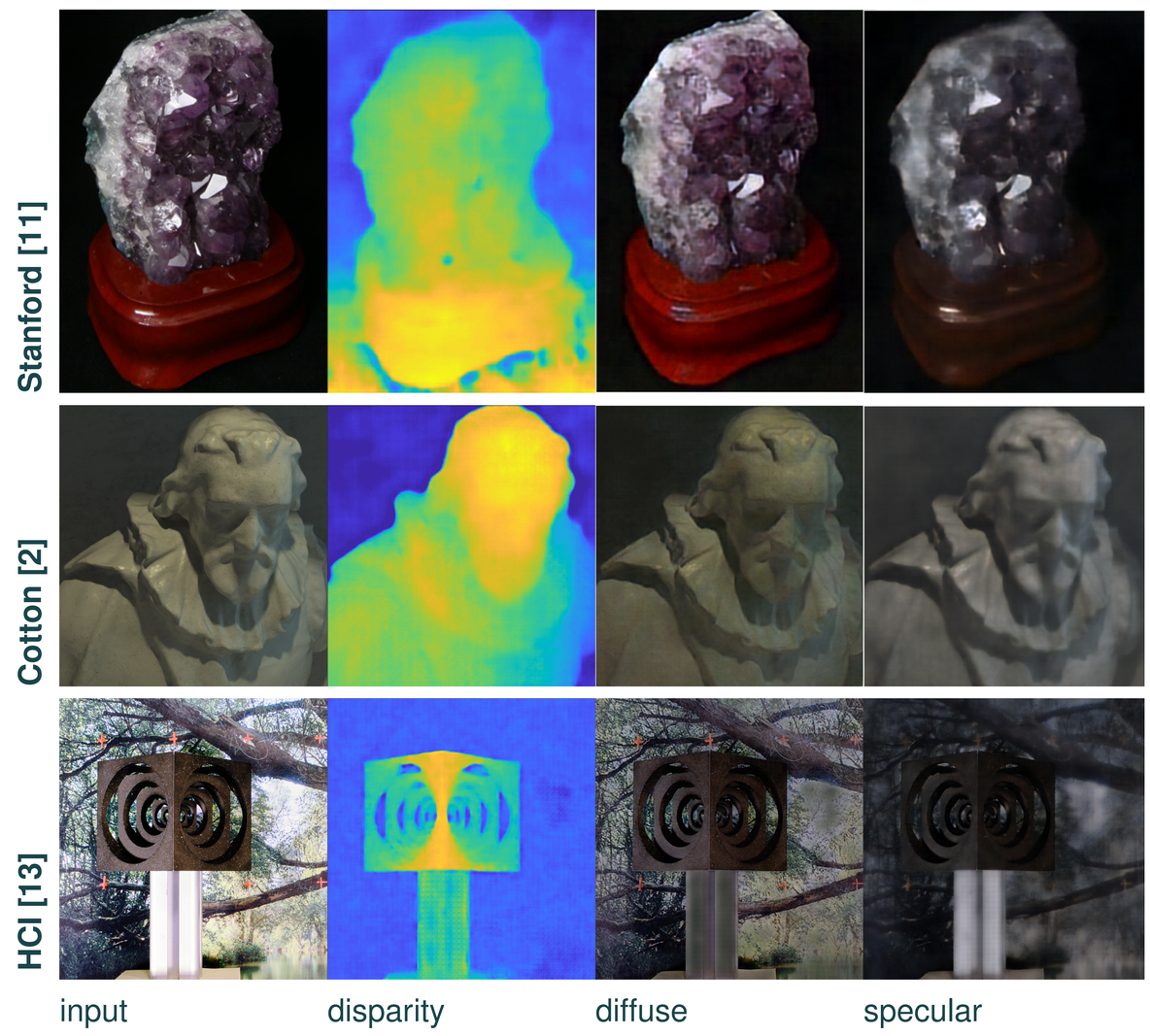
| | LMSE ×100 | | GMSE ×100 | | SSIM ×100 | |
|---|---|---|---|---|---|---|
| | diff. | spec. | diff. | spec. | diff. | spec. |
| **Ours** | 0.15 | **0.11** | 0.28 | **0.23** | **80.08** | **81.37** |
| Alperovich [1] | **0.12** | 0.45 | **0.22** | 1.04 | 74.98 | 48.46 |
| Sulc et al. [9] | **0.12** | 0.47 | 0.24 | 1.01 | 75.43 | 47.25 |
| Shi et al. [7] | 0.34 | 0.15 | 0.5 | 0.39 | 63.02 | 73.71 |

## Results for the real-world and synthetic light fields

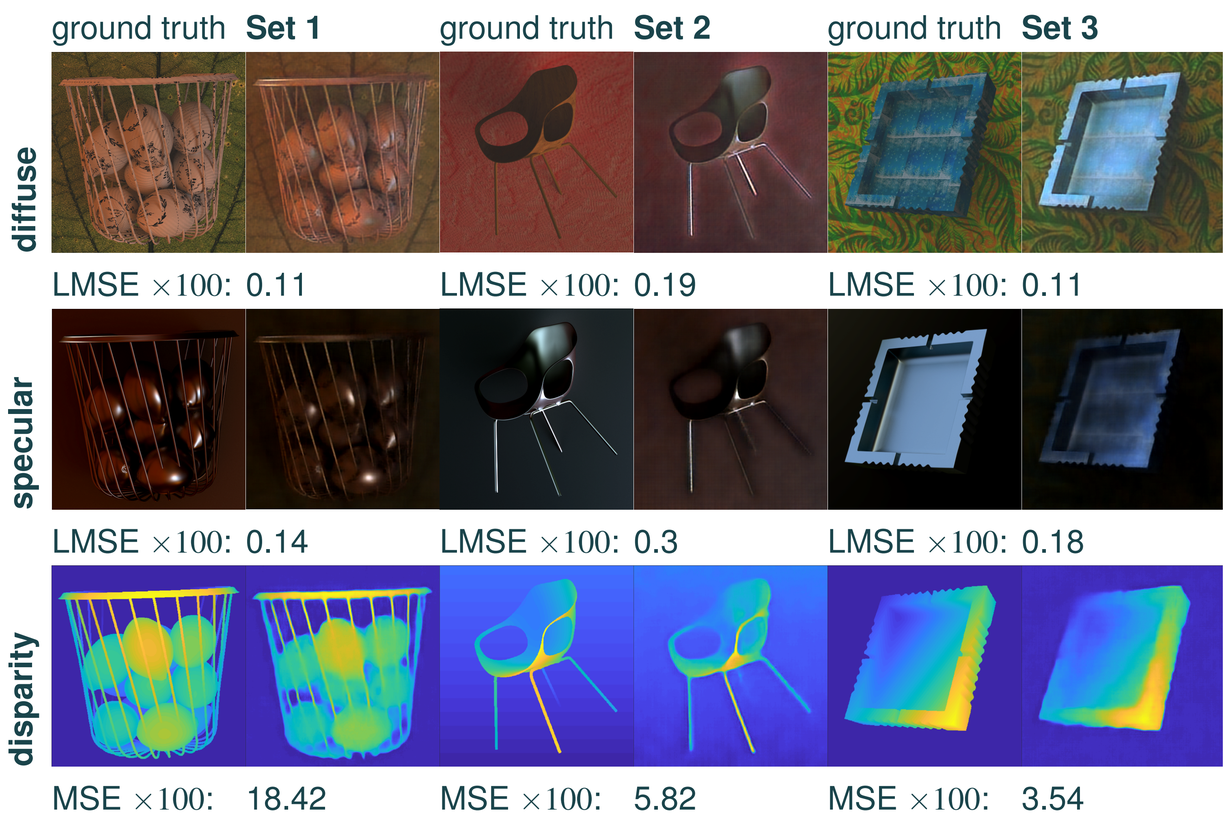
### Lytro Illum plenoptic camera




### various sources




## Visual and quantitative evaluations




## References

[1] A. Alperovich, O. Johannsen, M. Strecke, and B. Goldluecke. Shadow and specularity priors for intrinsic light field decomposition. In *EMMCVPR*, 2017.
[2] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke. A dataset and evaluation methodology for depth estimation on 4D light fields. In *ACCV*, 2016.
[3] H. Jeon, J. Park, G. Choe, J Park, Y. Bok, Y. Tai, and I. Kweon. Accurate depth map estimation from a lenslet light field camera. In *CVPR*, 2015.
[4] O. Johannsen, A. Sulc, and B. Goldluecke. What sparse light field coding reveals about scene structure. In *CVPR*, 2016.
[5] M. Levoy. Light fields and computational imaging. *Computer*, 39(8):46–55, 2006.
[6] Steven Shafer. Using color to separate reflection components. *Color Research & Application*, 10(4):210–218, 1985.
[7] J. Shi, Y. Dong, H. Su, and S. Yu. Learning non-lambertian object intrinsics across shapenet categories. In *CVPR*, 2017.
[8] P. Srinivasan, T. Wang, A. Sreelal, R. Ramamoorthi, and R. Ng. Learning to synthesize a 4D RGBD light field from a single image. In *ICCV*, 2017.
[9] A. Sulc, A. Alperovich, N. Marniok, and B. Goldluecke. Reflection separation in light fields based on sparse coding and specular flow. In *VMV*, 2016.
[10] M. Tao, J. C. Su, T. C. Wang, J. Malik, and R. Ramamoorthi. Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras. *TPAMI*, 38(6):1155–1169, 2015.
[11] V. Vaish and A. Adams. The (New) Stanford Light Field Archive. http://lightfield.stanford.edu, 2008.
[12] S. Wanner and B. Goldluecke. Globally consistent depth labeling of 4D light fields. In *CVPR*, pages 41–48, 2012.
[13] S. Wanner, S. Meister, and B. Goldluecke. Datasets and benchmarks for densely sampled 4D light fields. In *VMV*, 2013.