UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

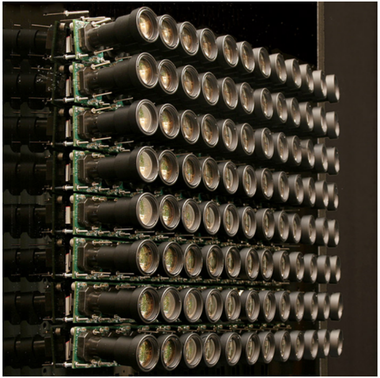
Visual Learning Lab Heidelberg

# Learning to Think Outside the Box:
# Wide-Baseline Light Field Depth Estimation with EPI-Shift

Titus Leistner

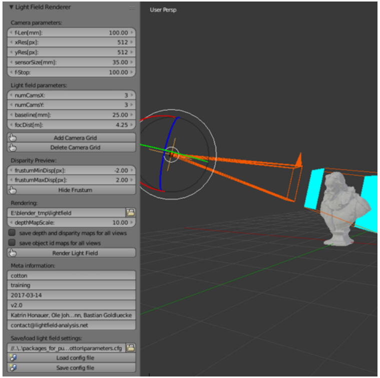
Hendrik Schilling, Radek Mackowiak, Stefan Gumhold, Carsten Rother

# Light Field Photography



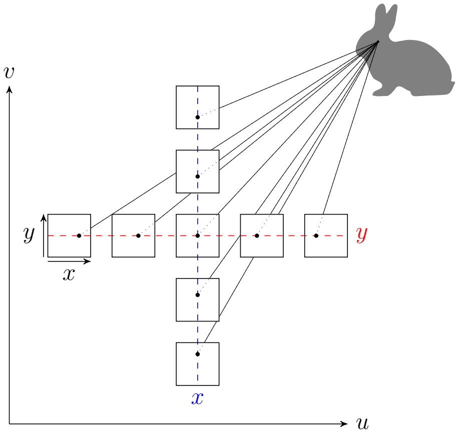
Camera array


Lytro camera



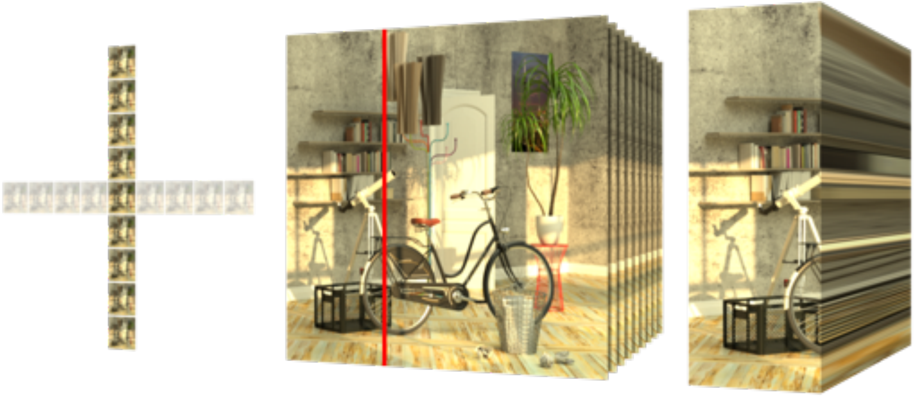

Rendering

# Disparity Estimation




camera indices $(u, v)$, image coordinates $(x, y)$

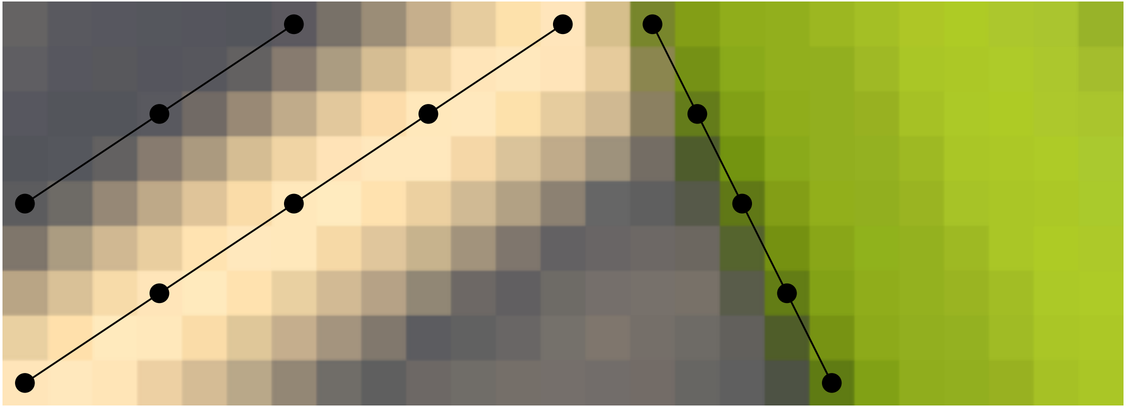
# EPI Extraction from a Cross Light Field



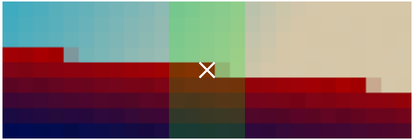
# EPI Extraction from a Cross Light Field

Epipolar Plane Image (EPI)

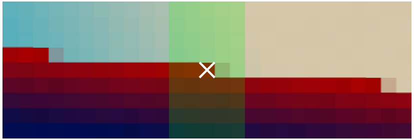
# Naive Disparity Estimation



# Basic Idea of EPI-Shift

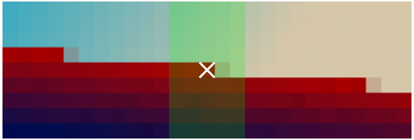


$d_1 = 10\text{px}$



$d_2 = 11\text{px}$

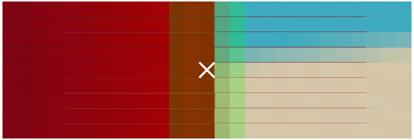
# Basic Idea of EPI-Shift



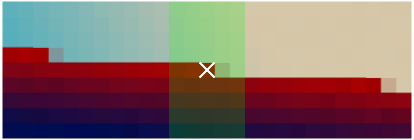
$d_1 = 10\text{px}$

EPI-Shift

by 10px



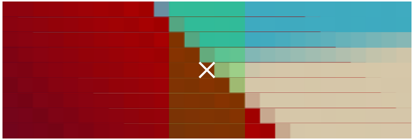
$d'_1 = 0\text{px}$



$d_2 = 11\text{px}$

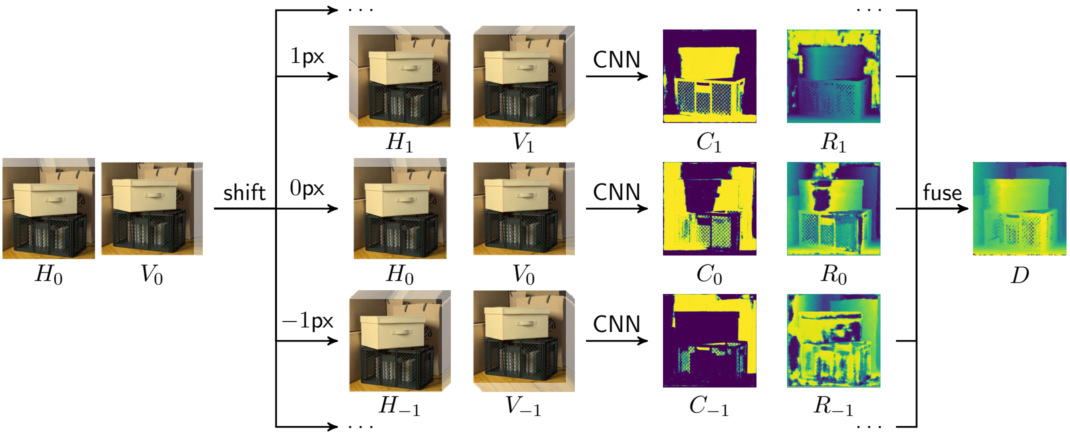
EPI-Shift

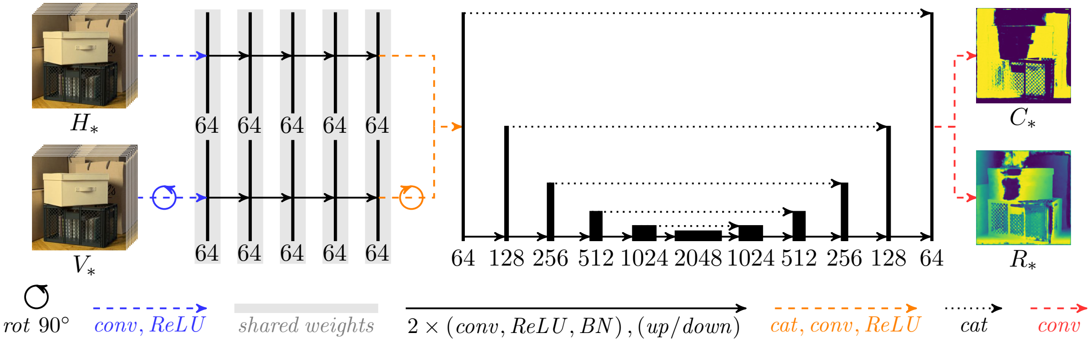
by 10px



$d'_2 = 1\text{px}$

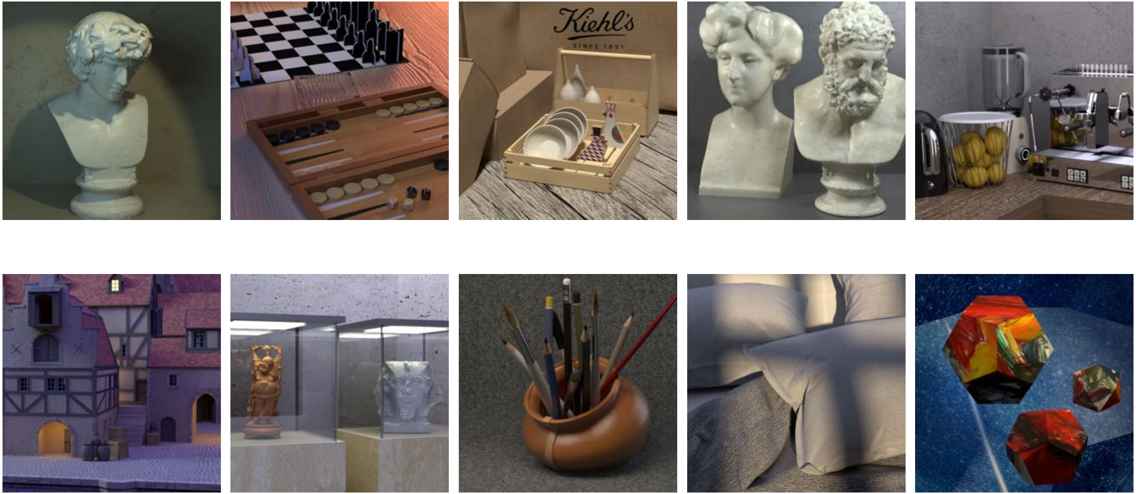
## EPI-Shift Pipeline



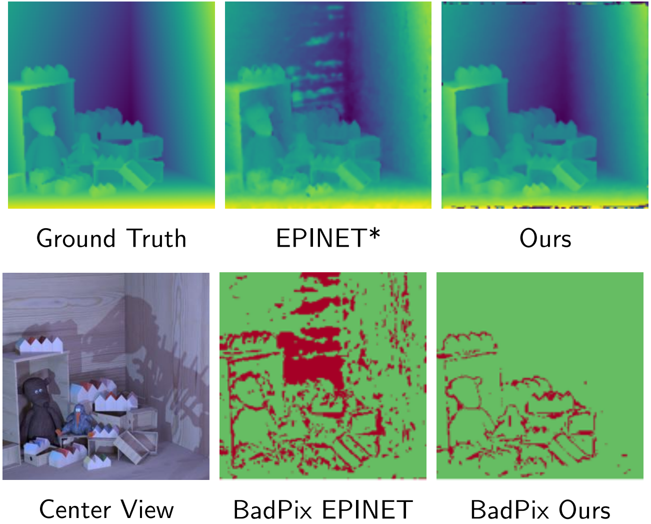


$H_*$
$V_*$
64 64 64 64 64
64 64 64 64 64
64 128 256 512 1024 2048 1024 512 256 128 64
$C_*$
$R_*$
rot 90°
conv, ReLU
shared weights
2 × (conv, ReLU, BN), (up/down)
cat, conv, ReLU
cat
conv

# Training Data

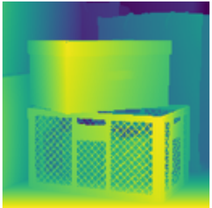


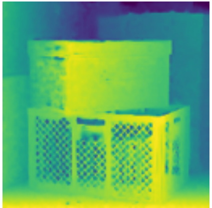
# Results

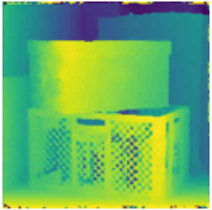



* C. Shin, H.-G. Jeon, Y. Yoon, I. S. Kweon, and S. J. Kim. EPINET: A Fully-Convolutional Neural Network Using Epipolar Geometry for Depth from Light Field Images. In CVPR, 2018

# Results

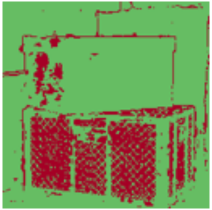

Ground Truth


EPINET


Ours

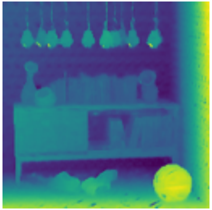

Center View


BadPix EPINET


BadPix Ours

# Results

Ground Truth



EPINET

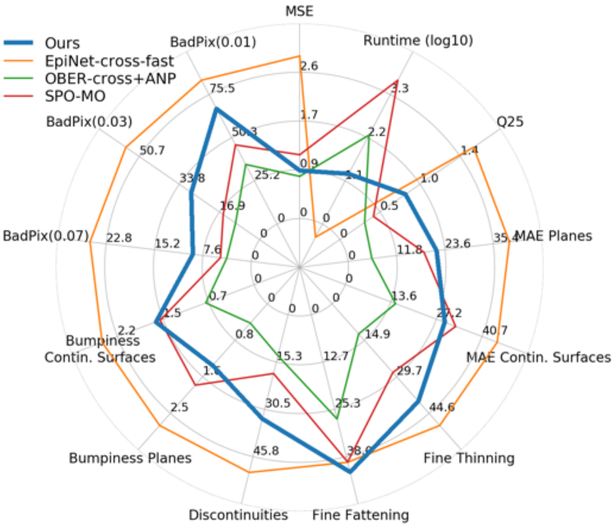
Ours

Center View

BadPix EPINET

BadPix Ours

# Results

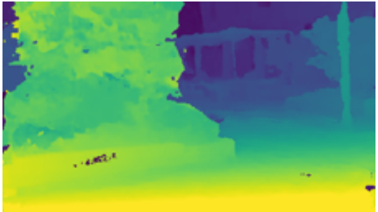


Ours
EpiNet-cross-fast
OBER-cross+ANP
SPO-MO
MSE
Runtime (log10)
Q25
MAE Planes
MAE Contin. Surfaces
Fine Thinning
Fine Fattening
Discontinuities
Bumpiness Planes
Bumpiness Contin. Surfaces
BadPix(0.07)
BadPix(0.03)
BadPix(0.01)

# Results

Center View

EPINET



Ours

## Contribution

- EPI-Shift for disparity range invariance
- U-Net for long-range smoothness
- State-of-the-art benchmark results

# Poster 108

# References I


[1] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991.

[2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. Patchmatch: A randomized correspondence algorithm for structural image editing. *ACMTOG*, 28(3):24:1–24:11, 2009.

[3] R. C. Bolles, H. H. Baker, and D. H. Marimont. Epipolar-plane image analysis: An approach to determining structure from motion. *International Journal of Computer Vision*, 1(1):7–55, Mar 1987.

[4] E. Brachmann, A. Krull, S. Nowozin, J. Shotton, F. Michel, S. Gumhold, and C. Rother. Dsac - differentiable ransac for camera localization. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[5] W. Commons. Lytro light field camera front, 2013. File:Lytro Light Field Camera-front oblique-fs PNr $^\circ$0405.jpg.

[6] M. Faraday. Thoughts on ray vibrations. *Philosophical Magazine*, 28(188), May 1846.

[7] S. Heber and T. Pock. Shape from light field meets robust pca. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 751–767, Cham, 2014. Springer International Publishing.

[8] S. Heber and T. Pock. Convolutional networks for shape from light field. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3746–3754, June 2016.

[9] S. Heber, W. Yu, and T. Pock. *U-shaped Networks for Shape from Light Field*. 9 2016.

[10] S. Heber, W. Yu, and T. Pock. Neural epi-volume networks for shape from light field. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2271–2279, Oct 2017.

[11] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke. A dataset and evaluation methodology for depth estimation on 4d light fields. In *Asian Conference on Computer Vision*. Springer, 2016.

[12] H. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y. W. Tai, and I. S. Kweon. Depth from a light field image with learning-based matching costs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2018.

# References II


[13] H.-G. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y.-W. Tai, and I. S. Kweon. Accurate depth map estimation from a lenslet light field camera. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1547–1555. IEEE, 2015.

[14] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The Digital Michelangelo Project: 3D scanning of large statues. In *Proceedings of ACM SIGGRAPH 2000*, pages 131–144, July 2000.

[15] H. Lin, C. Chen, S. B. Kang, and J. Yu. Depth recovery from light field using focal stack symmetry, 12 2015.

[16] A. Neri, M. Carli, and F. Battisti. A multi-resolution approach to depth field estimation in dense image arrays, 09 2015.

[17] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. 2005.

[18] H. Schilling, M. Diebold, C. Rother, and B. Jähne. Trust your model: Light field depth estimation with inline occlusion handling. In *CVPR*, 2018.

[19] H. Sheng, P. Zhao, S. Zhang, J. Zhang, and D. Yang. Occlusion-aware depth estimation for light field using multi-orientation epis. 74, 09 2017.

[20] C. Shin, H.-G. Jeon, Y. Yoon, I. S. Kweon, and S. J. Kim. Epinet: A fully-convolutional neural network using epipolar geometry for depth from light field images. In *CVPR*, 2018.

[21] S. Wanner and B. Goldlücke. Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans. Pattern Analysis Machine Intelligence*, 36:606–619, 2014.

[22] B. S. Wilburn, M. Smulski, H.-H. K. Lee, and M. A. Horowitz. Light field video camera. In *Media Processors 2002*, volume 4674, pages 29–37. International Society for Optics and Photonics, 2001.

[23] S. Zhang, H. Sheng, C. Li, J. Zhang, and Z. Xiong. Robust depth estimation for light field via spinning parallelogram operator. *Computer Vision and Image Understanding*, 145:148 – 159, 2016. Light Field for Computer Vision.