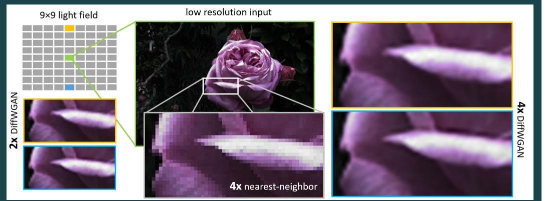
An Epipolar Volume Autoencoder with Adversarial Loss for Deep Light Field Super-Resolution

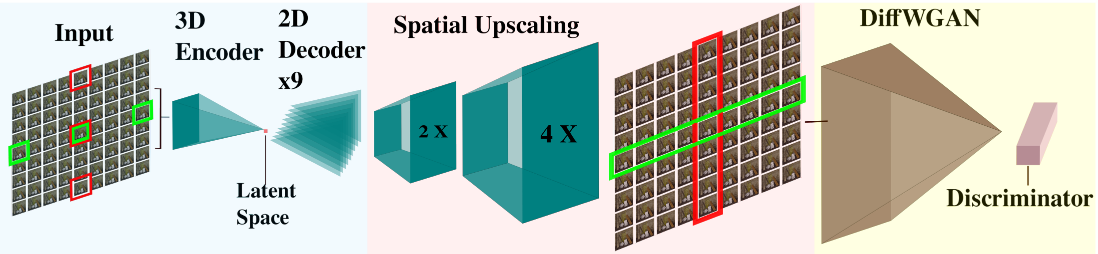
Minchen Zhu, Anna Alperovich, Ole Johannsen, Antonin Sulc and Bastian Goldluecke

Computer Vision and Image Analysis
University of Konstanz

Contributions

- a novel WGAN loss that penalizes the difference between angular and spatial derivatives of the generated light field and its ground truth
- in contrast to the classic GAN architectures for super-resolution [2] we added derivative information to the discriminator besides the ground truth image





Motivation

- Lightfields capture an additional **angular information**.
- The sampling of angular coordinates makes **trade-off** between spatial and angular resolution.
- Instead of hallucination like in 2D images, the lightfield profits from the **angular samples**.
- A **GAN autoencoder for lightfield superresolution** which is using the angular information.

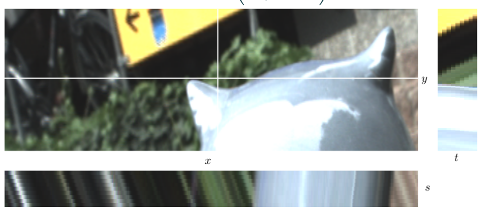
We need only three samples along each angualr components to increase:

- the **angular** resolution by a **factor of three**
- the **spatial** resolution by a **factor two or four**

Lightfields

A light field is defined on 4D ray-space $R = \Pi \times \Omega$, a ray is identified by four coordinates

$$\mathbf{r} = \left(\underbrace{s, t}_{\text{viewpoint}} \overbrace{y, x}^{\text{image}} \right) \tag{1}$$



Proposed Architecture

- **Input** 3 vertical and horizontal views are split into 48 $\times$ 48 patches with 16 overlapping pixels, in total $9 \times 48 \times 48$ input volume.
- **Latent space** The input volumes are downscaled spatially by encoder to $3 \times 3 \times 3$. The features are decoded and upscaled separately by **nine 2D decoders**.
- **Output** The decoded sub-aperture views are spatially upscaled twice to $9 \times 96 \times 96$.

Loss

A **high resolution** lightfield L_h and its **low-resolution** L_l counterpart. A following is passed to the input of the discriminator D

$$L = (L_h, \nabla L_h) \tag{2}$$

Where ∇ denotes a gradient along angular and spatial dimensions.
The total loss of generator $E_{\text{total}}(G)$ consist of two terms:

$$E_{\text{total}}(G) = E_{\text{rec}}(G) + E_{\text{WGAN}}(G) \tag{3}$$

The $E_{\text{rec}}(G)$ is a L^2 reconstruction loss between the high resolution output of the generator $G(L_l)$ and the ground truth L_h plus their derivatives:

$$E_{\text{rec}}(G) = (1 - \exp(-2L_h)) \|G(L_l) - L_h\| + \|\nabla G(L_l) - \nabla L_h\| \tag{4}$$

The $E_{\text{WGAN}}(G)$ is based on Wasserstein GAN

$$E_{\text{WGAN}}(G) = \overline{D(G(L_l))} \tag{5}$$

discriminator loss $E_{\text{WGAN}}(D)$:

$$E_{\text{WGAN}}(D) = \overline{D(L_h) - D(G(L_l))} \tag{6}$$

This work was supported by the ERC Starting Grant "Light Field Imaging and Analysis" (LIA 336978, FP7-2014).
Presented at the Conference on Computer Vision and Pattern Recognition (CVPR), Los Angeles, USA, July 2019.

An Epipolar Volume Autoencoder with Adversarial Loss for Deep Light Field Super-Resolution

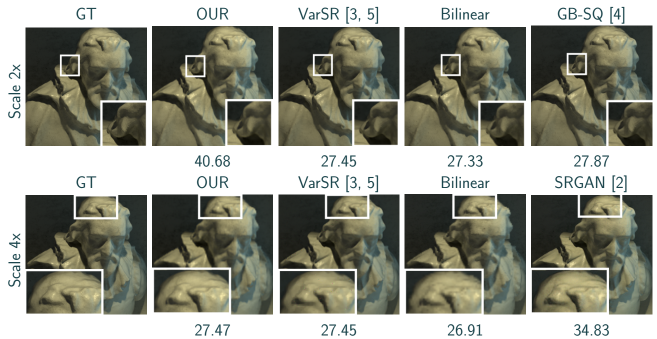
Minchen Zhu, Anna Alperovich, Ole Johannsen, Antonin Sulc and Bastian Goldluecke

Computer Vision and Image Analysis
University of Konstanz

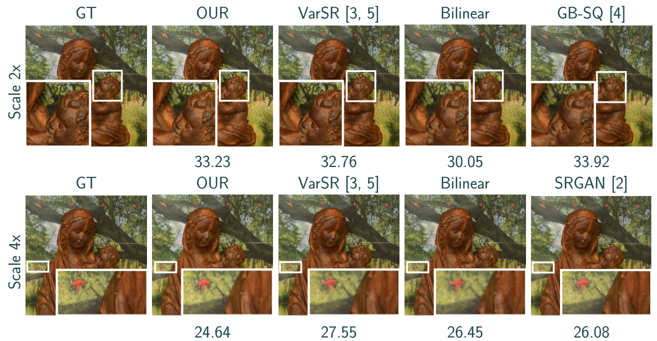
Experiments

Results

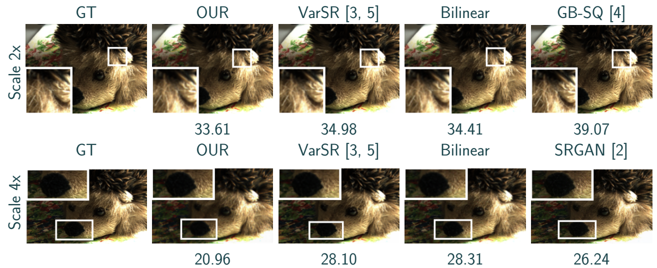
Synthetic Lightfield [1]



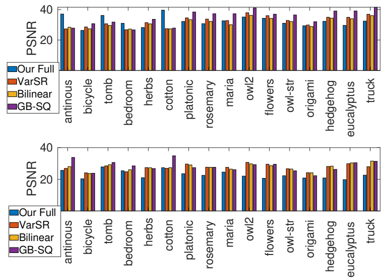
Gantry Lightfield [6]



Lytro Lightfield



Results



References

[1] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, A dataset and evaluation methodology for depth estimation on 4d light fields. In *Proc. ACCV*, 2016.

[2] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, Photo-realistic single image super-resolution using a generative adversarial network. *Proc. CVPR*, pages 105–114, 2017.

[3] S. Pujades, B. Goldluecke, and F. Devernay, Bayesian view synthesis and image-based rendering principles. In *Proc. CVPR*, 2014.

[4] M. Rossi and P. Frossard, Geometry-consistent light field super-resolution via graph-based regularization. *IEEE Transactions on Image Processing*, 27:4207–4218, 2018.

[5] S. Wanner and B. Goldluecke, Variational light field analysis for disparity estimation and super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(3):606–619, 2014.

[6] S. Wanner, S. Meister, and B. Goldluecke, Datasets and benchmarks for densely sampled 4D light fields. In *Vision, Modelling and Visualization (VMV)*, 2013.