Spatial and Angular Reconstruction of Light field Based on Deep Generative Networks

Nan Meng, Tianjiao Zeng, Edmund Y. Lam
Department of Electrical and Electronic Engineering, University of Hong Kong

Abstract

We propose a *generative network* based on high-dimensional convolution (HDC) layer for light field (LF) spatial and angular super-resolution (SR). Experiments are conducted on both synthetic and real-world LF scenes, and the proposed model outperforms existing methods in terms of PSNR and SSIM, and also generates more realistic spatial details with better fidelity.

Introduction

1. Considering the inherent geometry of LF data, we apply the HDC layer to involve information from both spatial and angular dimensions.
2. The proposed model *LFGAN* is then established by incorporating HDC layers to learn the high correlations among adjacent views.
3. A novel loss is defined to encourage *LFGAN* in recovering more realistic spatial details.

Problem formulation

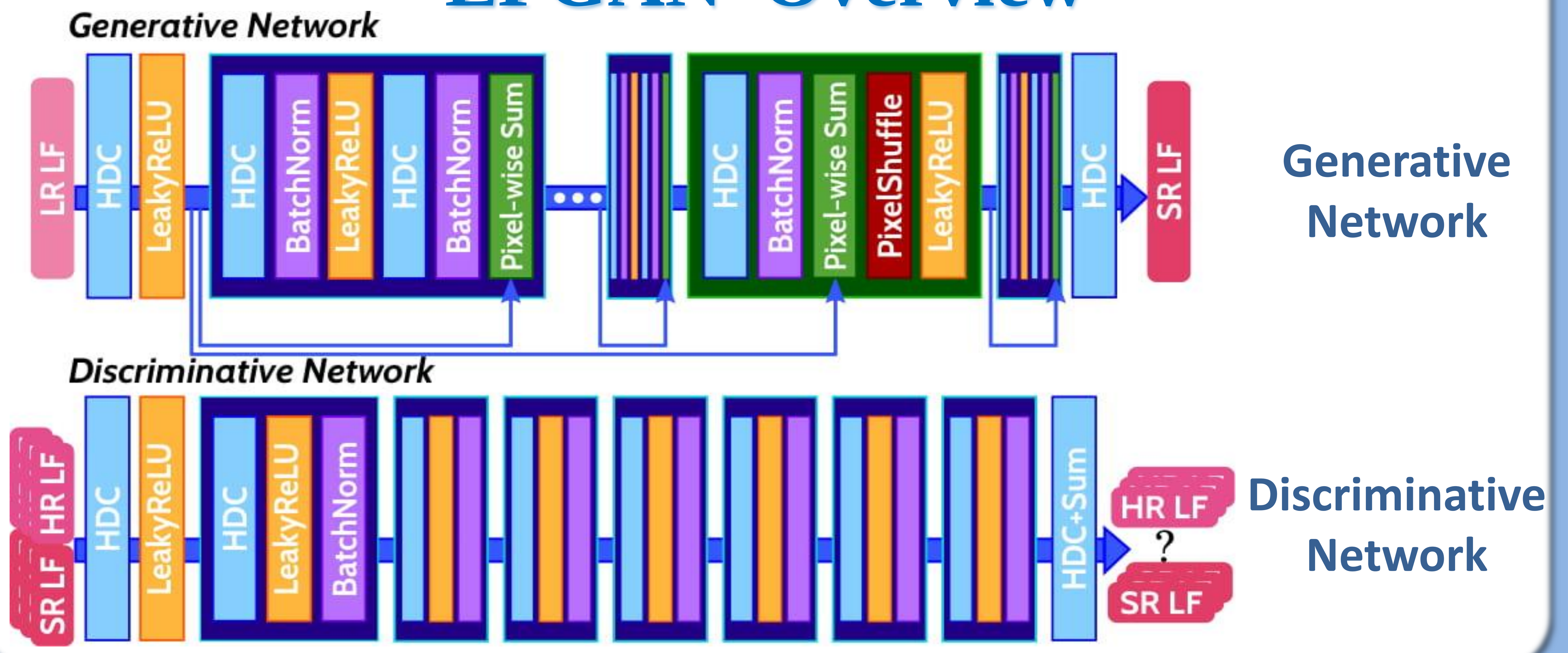
The LF reconstruction deals with the recovery of the high-resolution LF $I^H(x,y,u,v)$ from the corresponding low-resolution LF $I^L(x,y,u,v)$.

$$I^S(x,y,u,v) = g(I^L(x,y,u,v), \Theta),$$

where $g(\cdot)$ denotes the mapping formulated by the *generative network*, I^S is the super-resolved LF. Therefore, the restoration task is,

$$\Theta^* = \operatorname{argmin}\, \ell\big(I^H, g(I^L; \Theta)\big).$$

LFGAN Overview



Methods

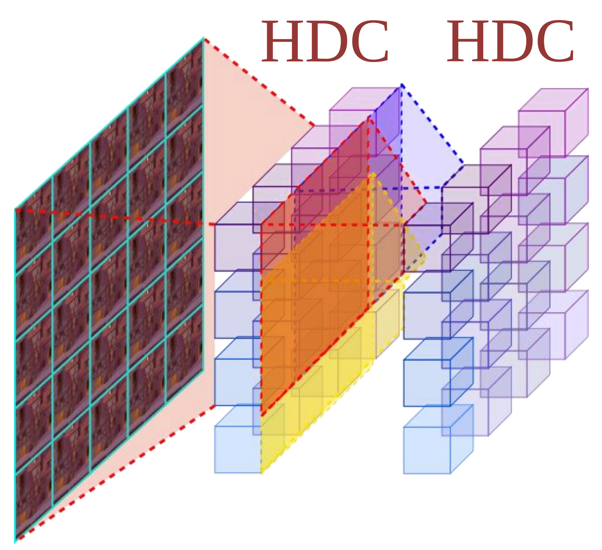


Fig.1 The angular receptive field of a 4D Residual Block.

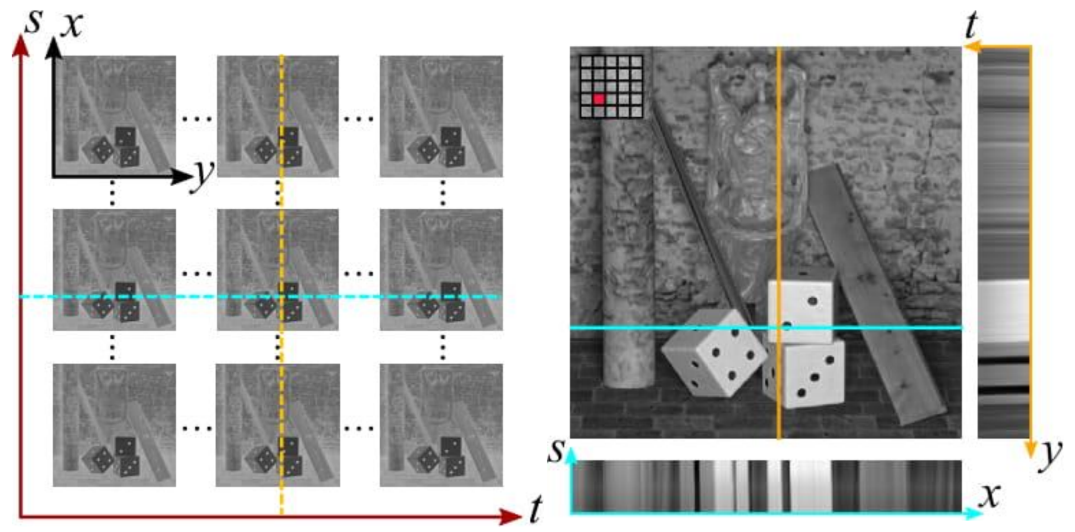


Fig.2 The learned 4D geometric features and EPI patterns.

Model Formulation

The training process can be formulated as solving the min-max problem.

$$\min_{\theta_G}\max_{\theta_D} \mathbb{E}_{I^H\sim\pi(I^H)}[\log D(I^H)] + \mathbb{E}_{I^L\sim\pi(I^L)}[\log(1 - D(G(I^L)))]$$

Loss function

We define a spatio-angular loss function to supervise the training process. Such loss is formulated as the weighted sum of a spatial content loss ℓ_S, an angular loss ℓ_A (based on MSE), and the adversarial loss ℓ_G.

$$\ell = \alpha \cdot \ell_S + \beta \cdot \ell_A + \gamma \cdot \ell_G,$$

where α, β, and γ are scalars.

- Spatial loss ℓ_S: $\ell_S = \frac{1}{ST}\sum_S\sum_T [f(I^H)_{s,t} - f(G(I^L))_{s,t}]$
- Angular loss ℓ_A: $\ell_A = \sum_Y\sum_T (E^{I^H}_{y,t}(x,s) - E^{I^S}_{y,t}(x,s))$
- Adversarial loss ℓ_G: $\ell_G = \sum_N \log[1 - D(G(I^L(x,y,s,t)))]$

Downsampling

The low-resolution LF I^L can be acquired by downsampling the high-resolution LF I^L according to the function:

$$I^L = \delta(B * I^H) + \eta,$$

where B is the Gaussian kernel with window size of 7 and standard deviation of 1.2. $\delta(\cdot)$ is the nearest neighbor downsampling operator, and η denotes the additive noise.

Results

Table 1. Quantitative evaluation on PSNR on synthetic LF and real-world LF for $\gamma_S = 4$. All numbers are measured in dB.

Algorithm	Buddha	Mona	Reflevtive20	Occlusions20
Bicubic	28.58	29.44	31.19	28.52
Yoon et al.	29.84	31.40	31.42	28.86
BM PCA+RR	30.43	32.68	33.07	30.45
LFNet	30.93	32.47	33.85	30.37
VDSR	30.48	31.39	32.32	29.84
MSLapSRN	30.98	32.74	32.43	30.85
RDN	30.99	31.80	33.86	31.46
LFGAN	**31.93**	**32.97**	**35.24**	**33.15**

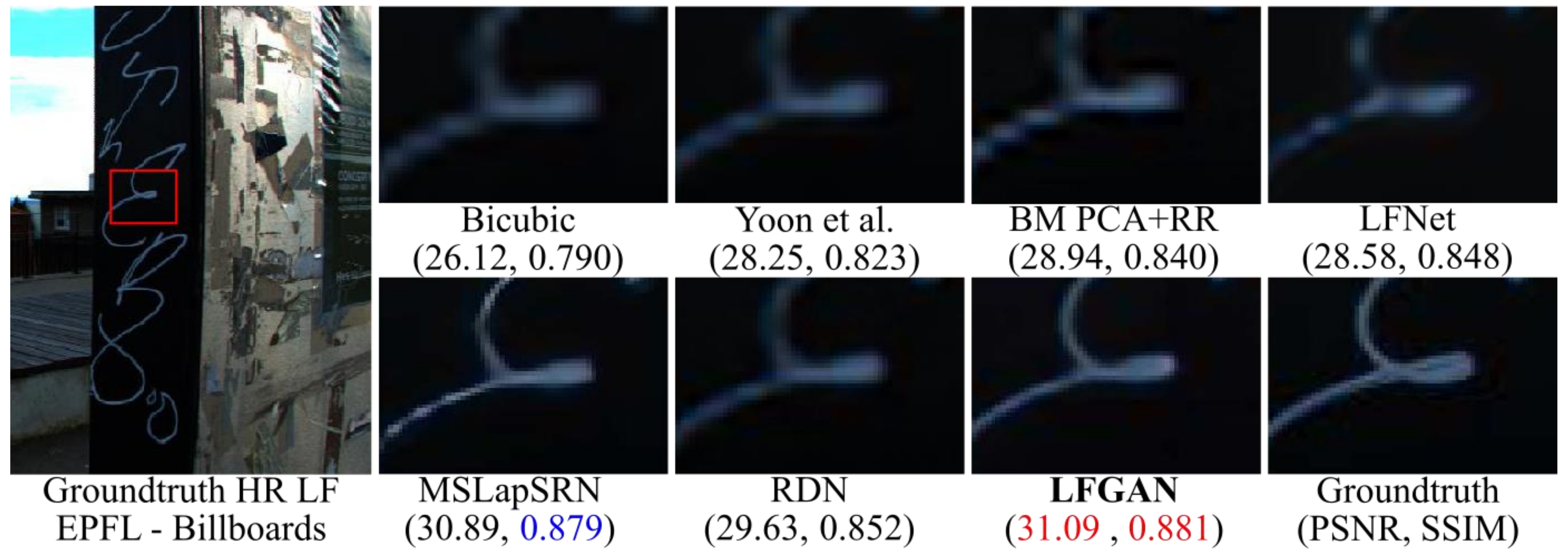


Fig 3. Visual and quantitative (PSNR, SSIM) comparisons on a real-world scene (EPFL-Billboards) for spatial $\times 4$ SR.

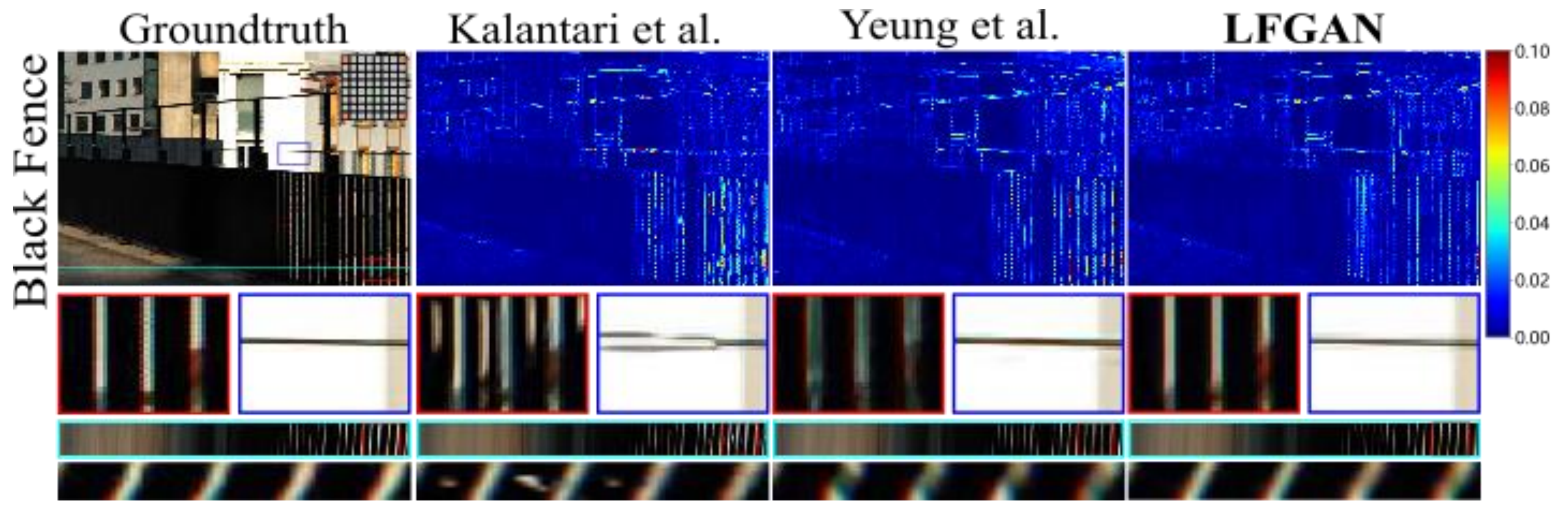


Fig 4. Visual comparisons on a real-world scene (EPFL-Black Fence) for angular $2 \times 2 \longrightarrow 8 \times 8$ SR.