Super-Resolving Light Fields

Shridevi Muthkhod
Department of Electrical Engineering
Stanford University
shridevi@stanford.edu

Abstract

With the availability of light field cameras like Lytro Illum, there is renewed interest in the field of light fields. One major hurdle in wide-spread adoption of this technology is the inherent trade-off in spatial and angular resolution of this camera which makes it either bulky or results in low resolution images in comparison with conventional cameras. Machine learning based data driven approaches have seen significant success in the area of image processing recently. Hence data driven techniques to super-resolve light field have been considered as a way to mitigate the trade-off between spatial and angular resolution. Current approaches either super-resolve the light field in the spatial or the angular domain. The goal of this paper is to discuss this trade-off while considering different machine learning techniques. Recent research in spatial and angular super-resolution methods is discussed in detail. A different implementation of angular super-resolution is also considered. It is seen that super-resolution only in the spatial domain has some challenges when reconstructing the light field. Hence angular information must be incorporated in the super-resolution strategy.

1. Introduction

Light field imaging involves representing the scene in a much richer format than conventional 2D imaging. A scene captured from a light field camera has not only the spatial information but also the depth information of the scene [1]. Computer vision applications like depth estimation, object classification, object recognition, etc. can largely benefit from light field images. There are a variety of cameras that capture light fields, ranging from the bulky arrays of conventional 2D cameras [2] to micro-lens array based cameras such as the Lytro Illum. But every light field camera inherently faces a trade-off between spatial and angular resolution where each camera lies at a different point on the trade-off curve. This is a major hurdle in wide-spread adoption of this technology resulting in either bulky cameras or low resolution images in comparison with conventional cameras.

Machine learning based data driven approaches have seen significant success in the area of image processing recently. Many different techniques have been proposed and used. Methods using deep CNNs have been developed in the last few years. One of the first methods is a three-layer CNN called Super-Resolution Convolutional Neural Network (SRCNN) [3] where C. Dong et al. demonstrate that the mapping from low resolution image to high resolution image can be cast as a CNN and also provide a model which is widely used as a reference even in light field super-resolution research. This network is therefore used in our comparisons.

One approach to light filed super-resolution is to use multiple shots with a programmable aperture that captures light field at the spatial resolution of the sensor in a compressed form. A good trade-off between spatial, angular and temporal resolution is proposed by Marwah et al. [4]. However, this approach requires modifications to the light field capturing device. Another way to mitigate the trade-off between spatial and angular resolution is by using data driven techniques to super-resolve light fields in the post processing step. Current approaches either super-resolve the light field in the spatial or the angular domain. The goal of this paper is to discuss this trade-off while considering different machine learning techniques. Recent research in spatial and angular super-resolution methods is discussed in detail. A different implementation of angular super-resolution is also considered. It is seen that super-resolution only in the spatial domain has some challenges when reconstructing the light field. Hence angular information must be incorporated in the super-resolution strategy.

In the following section, spatial super-resolution is discussed in detail. The effect of independent super-resolution on angular information is considered using the depth maps. Then angular super-resolution is discussed in the subsequent section. A different approach to angular super-resolution is also considered. The section on results and conclusions summarizes the key findings of the paper.

Figure 1. 7x7 sub-aperture image and depth map

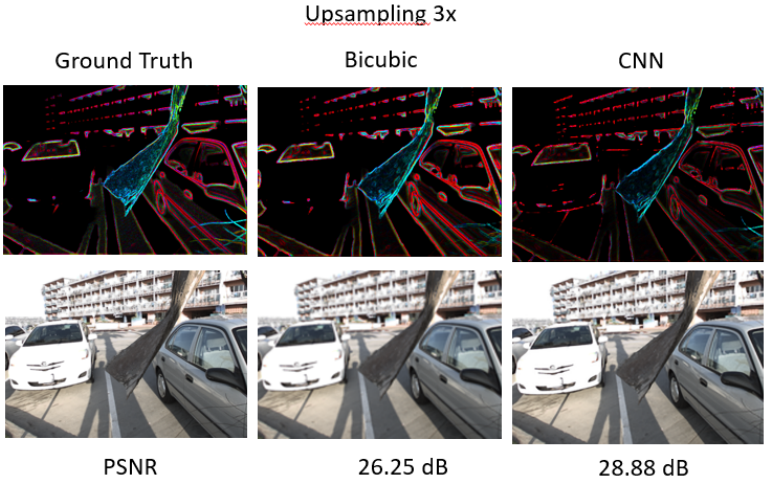

Figure 2. 7x7 sub-aperture image and depth map

2. Spatial Super-resolution

Spatial super-resolution considers super-resolving light field in the spatial domain. Traditional approaches assume a Lambertian surface and use estimated depth map to super-resolve the light field [5]. A Gaussian mixture model (GMM) is proposed instead in [6] to address spatial and angular super-resolution of light fields. These methods perform poorly on non-Lambertian surfaces, hence data driven approaches are considered.

2.1. Related Work

Yoon et al. [7] use a SRCNN to spatially super-resolve the light fields before feeding it to a angular super-resolving network. [8] looks at the entire stack of the sub-aperture images and super-resolves it. But this approach needs further processing and is able to achieve only 0.23 dB improvement.

2.2. Analysis

In this paper, therefore SRCNN is used to super-resolve the light field in the spatial domain. First the light field is down-sampled to obtain a low resolution light field. This light field is then up-sampled to the desired factor using bicubic interpolation. The up-sampled image is fed to the SRCNN to obtain the super-resolved light field. Depth maps are constructed for both bicubic interpolated and SRCNN super-resolved light field. Up-sampling factors of 2,3 and 4 are considered and the results are as shown in figure

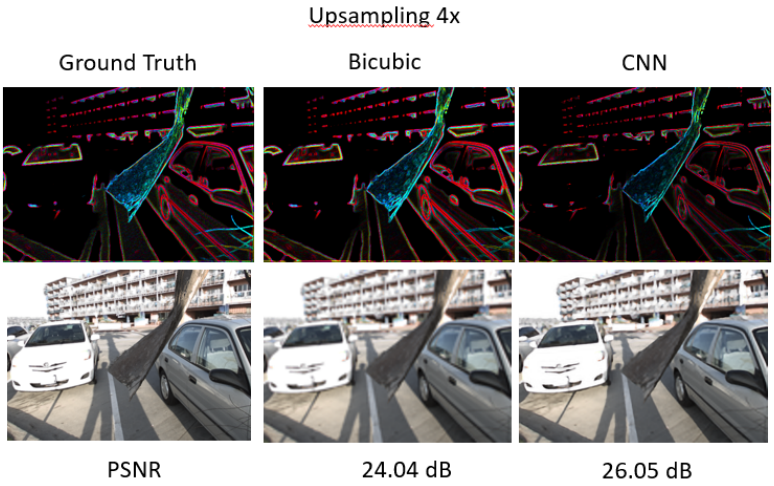

Figure 3. 7x7 sub-aperture image and depth map

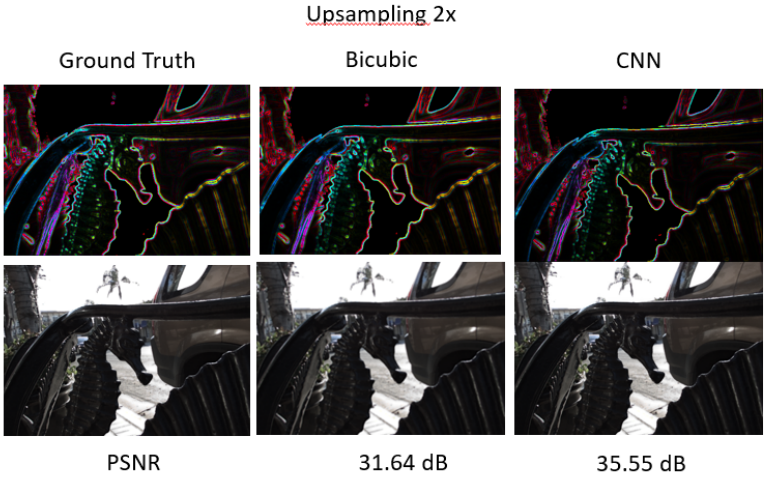

Figure 4. 7x7 sub-aperture image and depth map

2. The angular information is distorted as up-sampling factor increases by following such an approach, as can be seen from the depth maps.

3. Angular Super-resolution

Super-resolution in angular domain seems to be a better approach to super-resolving light fields. Angular super-resolution considers super-resolving light fields in the angular domain. There are different ways of capturing the angular information in a light field. In a Lytro camera, each micro-lens captures one particular pixel from different angular views. This determines the angular resolution of the light field image. In order to enhance this resolution further without increasing the camera size, angular super-resolution can be achieved through data-driven approaches.

3.1. Related Work

[9] uses a CNN to learn the depth disparity map and another CNN to learn the color estimation. But the algorithm assumes Lambertian regions and hence fails when super-resolving reflective surfaces. [7] uses 3 CNNs to find one view in between two horizontal views, one between two vertical views and one centre view using all four sub-aperture images. Angular up-sampling by a factor greater than two is not possible using their method. They also use spatial super-resolution as described above, to super-resolve each sub-aperture image before feeding to the angular super-resolution CNN which might adversely affect the angular

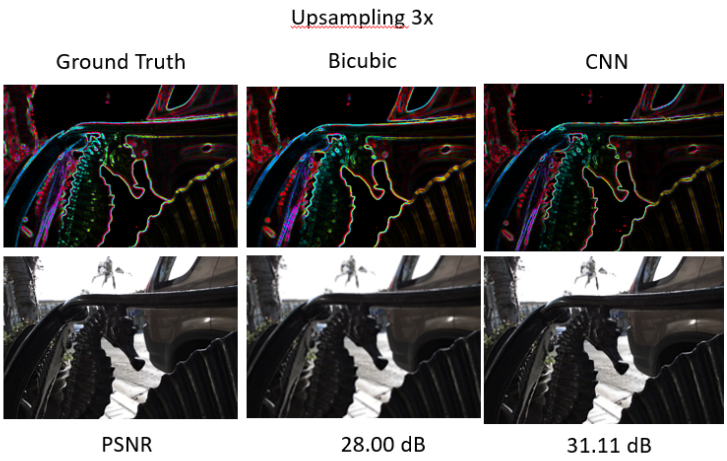

Figure 5. 7x7 sub-aperture image and depth map

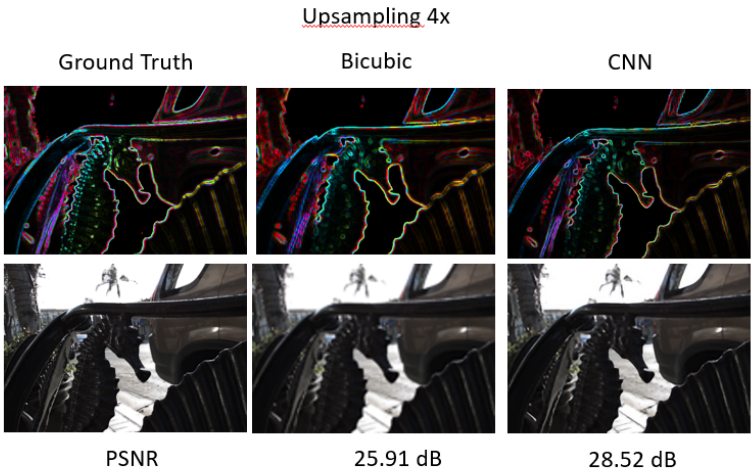

Figure 6. 7x7 sub-aperture image and depth map

information in the light field.

3.2. Analysis

In this paper, the problem of angular super-resolution using only the angular views is considered. The 2D angular image of all pixels is extracted from the light field. Super-resolved angular views are obtained using a similar network as SRCNN. The figure in 7 shows the implementation steps. Up-sampling factors of 2, 3 and 4 have been tested to obtain PSNR improvements from 1-3 dB in comparison with a simple bipolar interpolation. 8 shows a patch and compares this implementation with previous approaches at an up-sampling factor of 4. Yoon et al. [7] perform a spatial super-resolution with a factor 2, followed by angular super-resolution by a factor of 2. Hence for comparison, a similar spatial super-resolution followed by angular super-resolution is performed. The results are compared in 9 with a 5x5 real world image super-resolved to 10x10 angular views while also super-resolving the spatial dimension by a factor of 2.

4. Conclusions

The trade-off between spatial and angular resolution in light field cameras is discussed. Using data driven techniques to super-resolve light field in the post processing step thus reducing the cost and/ or size of the camera hardware is considered. Current approaches either super-resolve the light field in spatial or angular domain. This paper discusses

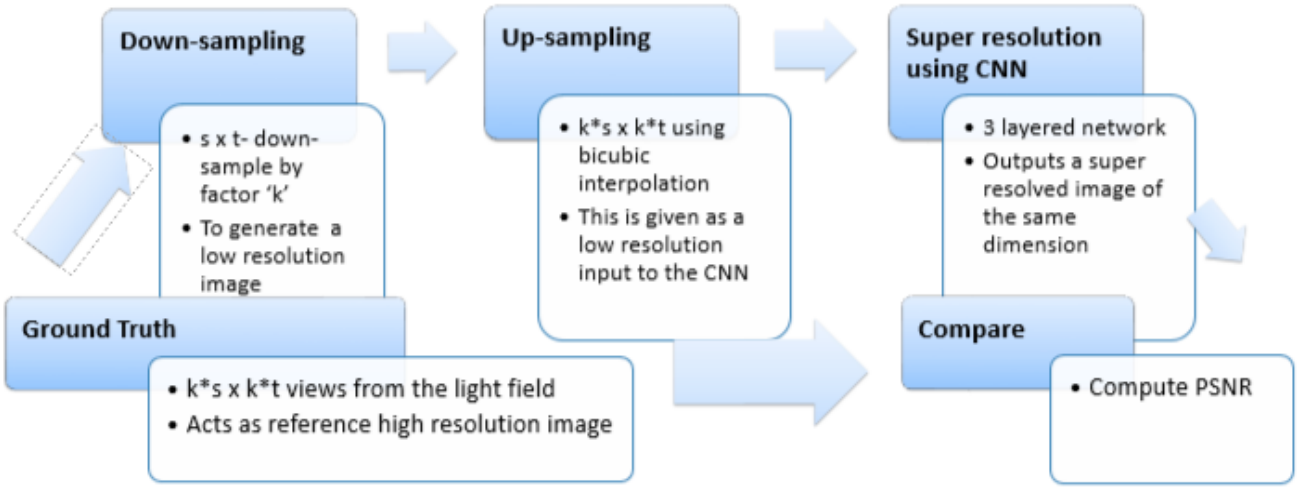

Figure 7. Block diagram of the angular super resolution implementation

Figure 8. Zoomed in views with angular up-sampling factor 4

this trade-off while considering different machine learning techniques. A different implementation of angular super-resolution is also considered. It is seen that super-resolution only in the spatial domain has some challenges when reconstructing the light field. Hence angular information must be incorporated in the super-resolution strategy.

5. Acknowledgements

The author would like to thank Dr. Donald Dansereau for his kind mentorship; and Prof. Gordon Wetzstein for providing the opportunity to work on this project.

References

[1] Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light Field Photography with a Hand-Held Plenoptic Camera. Technical report, April 2005.

[2] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. High performance imaging using large camera arrays. *ACM Trans. Graph.*, 24(3):765–776, July 2005.

[3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 184–199, Cham, 2014. Springer International Publishing.

[4] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive Light Field Photography using Overcomplete Dictionaries and Optimized Projections. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 32(4):1–11, 2013.

[5] Anat Levin, William T. Freeman, and Frédo Durand. Understanding camera trade-offs through a bayesian analysis of light field projections. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision – ECCV 2008*, pages 88–101, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[6] K. Mitra and A. Veeraraghavan. Light field denoising, light field superresolution and stereo camera based refocussing using a gmm light field patch prior. In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 22–28, June 2012.

[7] Y. Yoon, H. G. Jeon, D. Yoo, J. Y. Lee, and I. S. Kweon. Light-field image super-resolution using convolutional neural net-

Figure 9. Zoomed in views with spatial and angular up-sampling factors of 2,2

work. *IEEE Signal Processing Letters*, 24(6):848–852, June 2017.

[8] R. A. Farrugia, C. Galea, and C. Guillemot. Super resolution of light field images using linear subspace projection of patch-volumes. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):1058–1071, Oct 2017.

[9] Nima Khademi Kalantari, Ting-Chun Wang, and Ravi Ramamoorthi. Learning-based view synthesis for light field cameras. *ACM Trans. Graph.*, 35(6):193:1–193:10, November 2016.