Learning-Based Superresolution for 4D Light Field Images

Seung-Jae Lee and In Kyu Park
Department of Information and Communication Engineering, Inha University
Incheon 402-751, Korea
{*creative.sjlee@gmail.com, pik@inha.ac.kr*}

Abstract

A 4D light field image is represented in traditional 2D spatial domain and additional 2D angular domain. The 4D light field has a resolution limitation both in spatial and angular domains since 4D signals are captured by 2D CCD sensor. In this paper, we propose a dictionary learning-based superresolution algorithm in 4D light field domain to overcome the resolution limitation. The proposed algorithm improves both spatial and angular resolution by a factor of two. Experimental result shows that the proposed method outperforms the traditional method for the test images captured by a commercial light field camera, i.e. Lytro.

Keywords: Superresolution, dictionary learning, light field, spatial domain, angular domain, Lytro

1. Introduction

Recently, 4D light field image attracts much interest due to the post capturing capability of refocusing, viewpoint change, and depth map reconstruction. However, since 4D light field image signal shares 2D sensor resolution, both spatial and angular resolutions are limited. For example, the commercial light field camera, *i.e.* Lytro, has 9×9 angular resolution and 380×380 spatial resolution, which are all captured by 11 megarays CMOS sensor.

In order to alleviate the problem, it is necessary to develop an efficient superresolution technique which can be applied to 4D light field images. In this paper, we propose a novel superresolution algorithm using dictionary learning of 4D light field patches. The proposed algorithm consists of two steps. First, we extract a massive set of 4D patches with high resolution (HR) and low resolution (LR) correspondence. Then, a dictionary is constructed using K-means clustering. Note that dictionary learning is done in an offline processing. Second, given an LR light field image and the learned dictionary, we reconstruct a 4D light field image in an online processing to have double resolution both in spatial and angular domains.

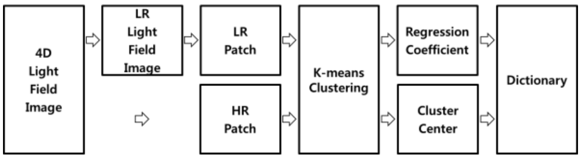


Fig.1. Overview of the dictionary training scheme.

2. Proposed Algorithm

2.1 Dictionary Training

The proposed algorithm collects a set of 4D light field images from real world scene captured by lenslet-based light field camera. However, it can be generalized to any type of light field cameras. We first downsample them to half resolution in both spatial and angular domains. The original and downsampled images are used for the training data. For each training data, we extract 4D patches with local pixels in spatial and angular domains. A 4D patch consists of a grid of spatial domain 2D patches which are extracted from the same location of each angular image.

In clustering step, we employ K-means clustering for easy and efficient clustering of patch samples. For each cluster, we estimate regression coefficients to generate the dictionary by solving the following linear least-squares problem:

$$\mathbf{C}^* = \underset{\mathbf{C}}{arg\min} \left\| \mathbf{H} - \mathbf{C}\begin{pmatrix}\mathbf{L}\\ \mathbf{1}\end{pmatrix} \right\|^2 \qquad (1)$$

where $\mathbf{C} \in \mathbb{R}^{n\times(m+1)}$ is the matrix of regression coefficients. n and m represent the number of pixels in a HR and LR patch, respectively. $\mathbf{H} \in \mathbb{R}^{n\times l}$ and $\mathbf{L} \in \mathbb{R}^{m\times l}$ are the matrix of HR and LR patches, where each column is the stack of pixels in a patch. l denotes the number of pairs of HR and LR patches in the same cluster. Fig. 1 shows the overview of the dictionary training scheme.

2.2 Reconstruction

In reconstruction step, we extract the 4D patches at all location of the input LR light field image. Then, the nearest cluster center in the dictionary is found for each 4D patch.

Fig. 2. Dataset captured by Lytro camera.

Let $\mathbf{C}^* \in \mathbb{R}^{n\times(m+1)}$ be the regression coefficient matrix of the corresponding nearest cluster center and $\mathbf{l} \in \mathbb{R}^m$ be the input LR patch in column form. Then, the 4D HR patch $\mathbf{h} \in \mathbb{R}^n$ is now reconstruct as a multiplication of input LR patches and the regression coefficients as follows.

$$\mathbf{h} = \mathbf{C}^* \binom{\mathbf{l}}{1} \qquad (2)$$

To complete 4D HR light field image, we compute (2) on each location and replace the LR patches with the reconstructed HR parches.

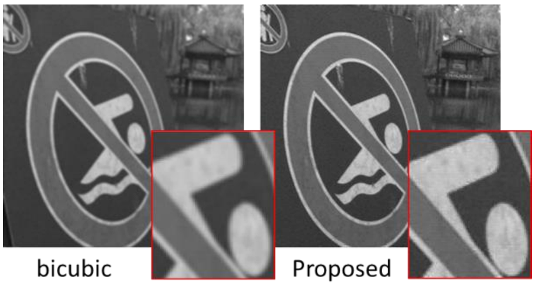
3. Experimental Result

We implement the proposed method on an Intel i7-3770K 3.5 GHz CPU with 16G RAM. We only apply superresolution to luminance component as other superresolution algorithms do. Test images are captured by Lytro. The captured 4D light field image has 360×360×8×8 resolution which is subsequently downsampled to half resolution (180×180×4×4). 4D HR and LR patches have 8×8×8×8 and 4×4×4×4 resolution, which are consequently in 4096×1 and 256×1 column vector form.

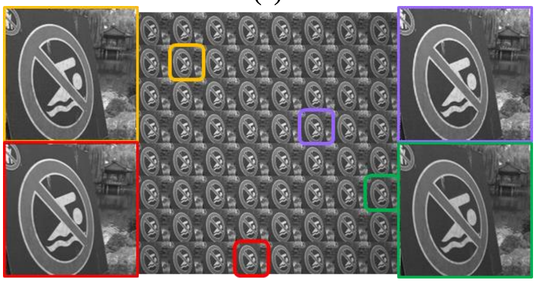
In our implementation, we randomly collect 200,000 patches from 40 4D light field images shown in Fig. 2. K is set 512 in K-means clustering. Proper K is determined under the trade-off relation between image quality and computation time.

The proposed dictionary learning-based superresolution method doubles the 4D resolution. Therefore, we obtain an 8×8×8×8 patch at each pixel location in LR image. 4D HR light field image is finally synthesized by composing all 4D HR patches.

In Fig. 3, we show the superresolution result with comparison to the conventional bicubic interpolation. It is shown that the proposed method outperforms both qualitatively and quantitatively. It reconstructs edges more clearly with higher PSNR. Note that the bicubic interpolation only increases spatial resolution, while the proposed method increases both spatial and angular resolution.



(a)



(b)

Fig. 3. Superresolution result. (a) Comparison with bicubic interpolation in spatial domain. PSNR is 31.15 dB for bicubic interpolation (left) and 32.72 dB for proposed method (right). (b) Result in angular domain. Angular images in even numbered rows and columns are the synthesized images.

4. Conclusion

This paper presented a novel dictionary-based superresolution method for 4D light field images. The proposed method improved the spatial resolution as well as angular resolution by factor of two. Experimental results demonstrated the improved resolution both quantitatively and qualitatively.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2013R1A2A2A01069181).

References

[1] C.-Y. Yang and M.-H. Yang, "Fast direct super-resolution by simple functions," Proc. ICCV, pp. 561-568, December 2013.

[2] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," Proc. CVPR, pp.1027-1034, June 2013.

[3] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Computer Science Technical Report (CSTR), 2005-2, April 2005.

[4] X. Huang and O. Cossairt, "Dictionary learning based color demosaicing for plenoptic cameras," Proc. CVPR Workshop, pp. 455-460, June 2014.