Light Field Super-Resolution: A Benchmark

we find that CNN-based single image SR without using any angular information (單純影像放大的方法)outperforms most light field SR methods(利用light field images放大的方法) even including learning-based ones.

因為通常光場放大的方法都是裡用sub-aperture之間的redundant information來增加spatial resolution

Light field super-resolution 的方法被分為三類：

1. projection-based: rely on the imaging principles of light field cameras.

* 最早由” J. Lim, H. Ok, B. Park, J. Kang, and S. Lee. Improving the spatial resolution based on 4d light field dat, 2009” 提出，利用2D sub-aperture之間pixel的shifting project在convex sets上可以加spatial resolution。
* “C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang. Learning a no-reference quality metric for single-image super resolution. 2017”利用refocusing principle和projected pixels from other views to the central view to get an all-in-focus image of the present scene.
* “C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. 2015”證明lenslet light field camera 保存頻率在spatial Nyquist rate上，並根據scene depth把light field samples project到target view

1. optimization-based：utilize various optimization frameworks to super-resolve light field images, relying on different mathematical or geometric modeling of the 4D light field structure.

* “T. E. Bishop and P. Favaro. The light field camera: Extended depth of field, aliasing, and superresolution, 2012”, 利用Lambertian reflectance and texture preserving priors在imaging model並重建HR利用variational Bayesian framework
* “K. Mitra and A. Veeraraghavan. Light field denoising, light field superresolution and stereo camera based refocussing using a gmm light field patch prior. 2012”, 假設每一張LF image之間的disparity map是一樣的，並用Guassian mixture model和linear minimum mean square error估算器，來補HR light field
* “S. Wanner and B. Goldluecke. Variational light field analysis for disparity estimation and super-resolution. 2014” 用EPIs來計算disparity map再藉由變異的framework來處理spatial 和angular
* “M. Rossi and P. Frossard. Graph-based light field super resolution. 2017” 耦合multiframe的方法和一個graph-based正則器增強LF結構，並避開計算disparity map。

1. learning- based
   * “R. A. Farrugia, C. Galea, and C. Guillemot. Super resolution of light field images using linear subspace projection of patch-volumes. 2017” 發現LF patch存在在低維子空間中，並用ridge regression學習一個LR到HR的linear mapping。
   * “Y. Yoon, H. G. Jeon, D. Yoo, J. Y. Lee, and I. S. Kweon. Learning a deep convolutional network for light-field image super-resolution. 2015” 第一個提出LF的SR 深度學習，用sub-aperture堆疊4-tuples然後餵進SRCNN「C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution.」

SR方法常常用不同datasets來比較，synthetic跟real-world image都應該納入考量，Data set使用：

HCI synthetic dataset：

10張[9\*9, 768\*768~1024\*720]

EPFL real-world dataset：

12張[9\*9, 625\*434], 為了解決vignetting的問題，用sub-aperture跟central view的平均intensity來校正。

Low resolution image也用不同降級假設：

不同的抽樣方法，即使用同樣的dataset，從HR->LR不同，會造成input不同無法公平比較，這裡使用兩種：

Bicubic(用MATLAB resize縮小2倍, 3倍),Gaussian(先對3\*3做高斯模糊,標準差設2)

與縮小2倍(周圍2\*2取平均),3倍(取3\*3中間pixel)的組合

選出4個LF SR方法根據：

近三年的state of art, 容易實現的code, 可調整抽樣方式

1. Projection-based, “C.-K. Liang and R. Ramamoorthi. A light transport framework for lenslet light field cameras. 2015” (利用” Occlusion aware depth estimation using light-field cameras. 2015”測量深度)：只產生central view的HR
2. Optimization-based, “M. Rossi and P. Frossard. Graph-based light field superresolution. 2017”
3. Learning-based, “R. A. Farrugia, C. Galea, and C. Guillemot. Super resolution of light field images using linear subspace projection of patch-volumes. 2017”,
4. “Y. Yoon, H. G. Jeon, D. Yoo, J. Y. Lee, and I. S. Kweon. Learning a deep convolutional network for light-field image super-resolution. 2015”, 此篇論文把裡面的SRCNN替代為VDSR。
5. 單純CNN放大 VDSR, ” J. Kim, J. Kwon Lee, and K. Mu Lee. Accurate image super resolution using very deep convolutional networks. 2016”

評量方法：

(直接視覺評估)PSNR, SSIM,

(perceptual quality ) VGG matric(R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang. The unreasonable effectiveness of deep features as a perceptual metric. In CVPR, 2018.)： 計算在VGG19 feature space中的誤差，

Ma’s matric(C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang. Learning a no-reference quality metric for single-image super resolution. 2017)

比較central view的SR：

Reconstruction accuracy：

Bicubic放大方法效果較差，因為他沒有考慮到視圖間的資訊，

* Non-learning based的方法在HCI dataset上表現較好，因為合成的圖有比較乾淨的內容。  
  Learning based的方法在EPFL上表現較好，因為可以處理比較多雜訊的real world image。

(Gaussian downsampling配2倍scale factor相當於做兩次的low-pass會產生更嚴重degraded的LR image，所以learning based的LFCNN可以較好的處理這種現象。)

(PSNR中VDSR比其他的方式好是因為他有大量的訓練資料)

Perceptual metric：

(用VGG metric比較時，反而是在PSNR表現不好的PRO表現不錯，可說明the tradeoff between reconstruction accuracy and perceptual quality)

Visual results：(用Gaussian downsampling with the scale factor of 3)

PRO在平滑區可以有不錯的放大效果，但邊緣效果不好

LFCNN在邊緣效果好，

GB介於PRO跟LFCNN之間

RR效果不好

VDSR有更好的邊緣效果

因為材質、平滑區比較難藉由訓練及學習，但可以藉由其他sub-aperture資訊取得，可以邊緣用DL平滑用projecting-based。

比較全部sub-aperture的SR：

EPFL的標準差比較大，因為現實照片比較會受到漸暈、硬體設備影響

LFCNN一般化：

原本該方法的論文中用k-folds的方法來訓練模型，但是一般來說不能使用自己來訓練自己，所以此篇用HCI訓練EPFL來測試，PSNR下降了約1.25，這是CNN-based的缺點it can be expected that LFCNN would benefit from increased training data in the same domain.

本文提出一些可以突破的點：

1. real-world LF data太少，但用litro 相機拍的又受限於spatial resolution，可以先用VDSR把LITRO拍的照片放大，再拿去CNN-based的SR方法train。
2. 不同光場相機或拍攝方式，作為訓練集會有明顯差異，為了解決domain shift的問題，可以考慮結合本light fields當作Graph來學習4D間的關聯性、或利用兩種domain 在CNN中相似的layer作為訓練集。