

A NOVEL DISPARITY-ASSISTED BLOCK MATCHING-BASED APPROACH FOR SUPER-RESOLUTION OF LIGHT FIELD IMAGES

Saber Farag and Vladan Velisavljevic

School of Computer Science and Technology, University of Bedfordshire, Luton, United Kingdom

ABSTRACT

Currently, available plenoptic imaging technology has limited resolution. That makes it challenging to use this technology in applications, where sharpness is essential, such as film industry. Previous attempts aimed at enhancing the spatial resolution of plenoptic light field (LF) images were based on block and patch matching inherited from classical image super-resolution, where multiple views were considered as separate frames. By contrast to these approaches, a novel super-resolution technique is proposed in this paper with a focus on exploiting estimated disparity information to reduce the matching area in the super-resolution process. We estimate the disparity information from the interpolated LR view point images (VPs). We denote our method as light field block matching super-resolution. We additionally combine our novel super-resolution method with directionally adaptive image interpolation from [1] to preserve sharpness of the high-resolution images. We prove a steady gain in the PSNR and SSIM quality of the super-resolved images for the resolution enhancement factor 8x8 as compared to the recent approaches and also to our previous work [2].

Index Terms — Resolution Enhancement, Light Field Image Super-Resolution, Block Matching, 4D Imaging

1. INTRODUCTION

Advances in technology of the plenoptic camera have brought important attention in both research community and industry to provide efficient scene refocusing at different depths without additional computational complexity [3, 4, 5, 6, 7]. This type of camera is based on the concept initially introduced by Adelson *et al.* [6] and further developed by Levoy *et al.* [3] by parameterizing the 4D LF. One of the common ways to represent a LF data is to consider it as a matrix of VPs, where each VP captures a 2D slice of the LF. In contrast to classical cameras, plenoptic cameras are capable of capturing images (2D LF slices) at multiple VPs at the same time owing to a system of multi-array lenses, a significant leap forward in the camera hardware technology. Conversely, presently available plenoptic camera technology is capable of providing only a limited depth-of-field and low spatial image resolution [8]. For example, the Lytro camera's final super resolution is restricted to only 300 x 300 pixels [9].

Several approaches have been proposed to enhance this limited spatial resolution. For instance, the authors in [7] and [8] used the EPIs to measure variational methods and disparity maps in order to calculate a LF image. Furthermore, in [10] and [11] the VP information is exploited for LF denoising application by extending the block matching 3D (BM3D) filter. In contrast to recent studies, and to the best of our knowledge we are the first to exploit BM to light-field super-resolution (LFSR) rather than motion estimation in classical video processing.

This study proposes a novel LFSR technique that is capable of enhancing spatial resolution of super-resolved LF images, while preserving low computational complexity. Our approach consists of two steps: first, we apply the image interpolation technique from [1] to each VP to enhance spatial resolution by the factor 4x4. This method is based on directionally adaptive wavelets and is able to preserve efficiently sharpness in the interpolated high-resolution images. Second, we estimate disparity information from these high-resolution images using the algorithm from [9]. We use then this information to facilitate block matching super-resolution by narrowing down the block search area in neighbouring VPs. Thus, the disparity pattern largely relies on the VPs of the scene from the plane of the LF. Furthermore, a LF block of dimension $n_a \times n_a \times k \times k$ that corresponds to a 4D block was considered as opposed to the BM3D, which searches and combines blocks only within a single image, whereas our light field block matching super-resolution (LFBMSR) approach operates across multiple views, exploiting similar blocks from different depths. Further, to obtain the final super resolved image 8x8 times, we use our recent LF hybrid super-resolution (LFHYSR) approach in [2]. Finally, we demonstrate the robustness of the proposed novel technique for solving LF processing tasks, namely the LFSR. As compared to the current LFSR images, the proposed novel approach efficiently produces far better quality images as opposed to the images obtained with the use of the MISR [12], and HYSISR [13].

The sections of the remaining paper are as follows. Section 2 contains the previous work of super-resolution for LF images. Section 3, shows our technical work, which explained the novel method we propose for BM to LF images and an overview of our recent method to enhance the resolution. Section 4 shows the experimental results and evaluation. Section 5 concludes the paper.

2. REVIEW OF RELATED WORK

Super-resolution of LF images could be divided into two major classes: variational optimization framework and projection-based algorithms. As an example of the variational optimization framework, in [14] the authors improved image features by outlining the variational Bayesian scheme and expanding the captured LF image's resolution. Thus, they evaluate both the LF and high-resolution (HR) of the depth map, although they manipulate prior knowledge about the scene, which helps to get extra data from accessible information. Their approach enhances image quality by recognizing the point spread function (PSF) of the plenoptic camera under Gaussian optics suppositions over numerous insight scenes, which suffers from limited spatial. The authors of [5] proposed a method by utilizing the EPI to figure variational techniques and continuous disparity maps to compute HR image. The authors of [8] performed Bayesian derivation to accomplish LFSR by applying Gaussian mixture model (GMM) over LF patches for LF image which uses a projection-based technique. In

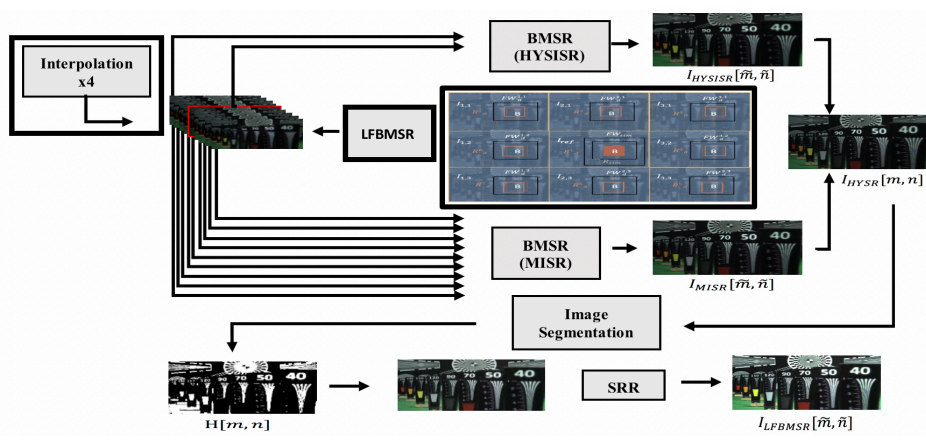


Figure 1. Proposed LF block matching super-resolution technique. Interpolation approach is used to enhance spatial resolution by 4x4 to preserve efficiently sharpness. LFBMSR is the process of narrowing down the block search area in neighboring VPs. The combination of MISR and HYSISR are performed using Laplacian Pyramid in order to develop a HYSISR image. In comparison, the LFBMSR outperforms both MISR and HYSISR in terms of the numbers and quality of the image generated.

their proposed technique, the evaluated camera with the current LFSR does not rely on its preset parameters. A simple method to light field block matching (LFBM) lies in applying recent block matching approaches to the VPs independently. Four articles [16, 17, 18, 19] are noted for the reader for reviews of the BM approaches applied to 2D and 3D images. The authors in [19] proposed method based on adaptive filtering, in order to improve the robustness of video sequence in real life. However, the proven determination upgrade factor of nonlocal strategies has for the most part been unassuming. Moreover, nonlocal procedures encounter piece coordinating challenges with vast relocations, rotational movement, and obscured edges. The EPIs technique is used to process the redundancies that exists between the VPs. The two advance approaches suggested in [20] which initially denoises EPIs, brought a certain spatial as well as an angular dimension (for instance sx -plane), and after that procedures, this initial measure utilizing the corresponding EPIs (i.e., yt -plane). Nonetheless, the above approaches merely take into consideration the 2D facets of the 4D LF. Despite the fact that these techniques misuse entirely the 4D structure of the LF, they neglect to bypass the video block matching 4D (VBM4D). For basic knowledge of BM3D, we recommend the work presented in [21] for readers because the image block matching methods discussed in the study have demonstrated high performance. The proposed LFBMSR in this paper is an extension of LFSR. We demonstrate that by altogether considering the 4D LF structure, we can fundamentally outperform the compete methods. In the following section we explain our proposed method. A block-scheme for the proposed LFBMSR method is presented in Figure 1.

3. BLOCK MATCHING SUPER-RESOLUTION

The LF is expressed as a 4D function¹. $\Omega \times \Pi \rightarrow \mathbb{R}, (x, y, s, t) \rightarrow L(x, y, s, t)$ where Ω denotes the light rays' spatial distribution, ordered by (x, y) , whereas Π represents their angular distribution, ordered by (s, t) . Hence, the LFBMSR is a novel approach added to our recent method in [2]. We have exploited in our method a corresponding two-step technique as BM3D. Firstly, we produce the first estimation (i.e., first iteration) in order to find the best similar block matching between all sub-aperture views; or viewpoint images (VPs). Secondly, we utilized the first estimate to implement the 4D transform domain and produce the estimated high-resolution image. We demonstrate the method implementation in detail. Firstly, a certain VP, referred to as the reference VP, is chosen as the central VP of LF. Secondly, we iterate over VPs to find similar blocks to the selected reference VP. Furthermore, the neighboring VPs which are found within the search window of size $n_a \times n_a$ are considered, and all these VPs are processed together. Thus, recursive iteration is carefully

performed over the 2D blocks of the reference VP. Furthermore, 4D blocks are generated through the combination of the 2D blocks of the reference VP and their corresponding neighboring VPs. Additionally, another recursive iteration is usually required in order to process a handful remaining pixels within the surrounding VPs. Finally, another reference VP will be chosen from the remaining VPs when the former reference VP and the surrounding neighboring VPs have been completely processed.

3.1 Blocks Formation

We denote by B the reference current block of size $k \times k$ in the reference VP I_{ref} , we create a 4D by exploiting redundancies within I_{ref} and its neighboring VPs $\{I_{s,t}\}, (s, t) \in [1, n_a] \times [1, n_a]$ 4D block is generated by searching all the neighboring VPs of the 2D block nearest to B that could be aligned to a disparity compensation step by applying block matching technique. Originally, we defined the 4D block dimension $n_a \times n_a \times k \times k$ as:

$$B_d(B) = \{R_d^{s,t} : R_d^{s,t} = \arg \min d(B, R_{s,t}), R_{s,t} \in FW_d^{s,t} : d(B, R_d^{s,t}) \leq \tau_d\} \quad (1)$$

with $d(B, R)$ representing the aligned quadratic distance which separates the different blocks, $FW_d^{s,t}$ denotes the search window over $I_{s,t}$ with size $n_a \times n_a$ focused on the location of B . Whereas τ_d represents the specified distance threshold for d over which set of blocks are usually considered similar. Consequently, the threshold τ_d aims to identify and reject certain blocks that are not satisfactorily similar in every respect to the selected reference block, and as a result, to be strong against occlusions. Figure 2, shows the search window of the BMSR algorithm outline.

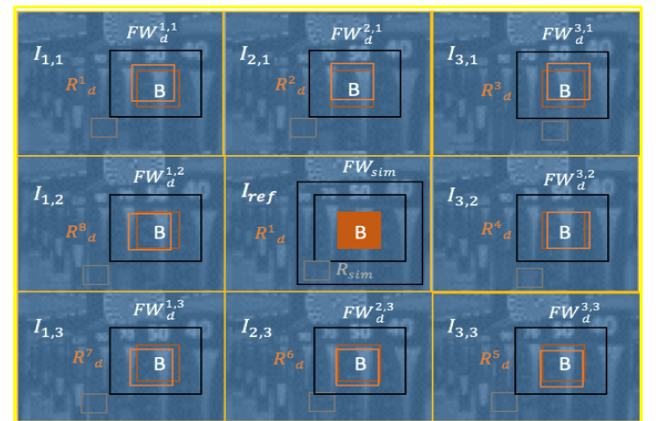


Figure 2. Search window of the BMSR algorithm outline. With the help of the search window, the VPs of each LF are processed in iteration by taking into consideration, the reference VP alongside its surrounding VPs. 4D blocks are usually acquired when VPs are processed with angular search window through the use of disparity compensated 2D blocks within the surrounding VPs with regard to a 2D block in the reference VP.

¹ Note that we neglect the chromatic and temporal dimensions and we restrict the VP to a 2D plane here.

3.2 Self-Similarities

The search is then done for the set of neighboring blocks similar to the reference selected VP block B , which is defined by:

$$\mathcal{B}_{sim}(B) = \{R_{sim} : d(B, R_{sim}) \leq \tau_{sim} R_{sim} \in FW_{sim}\} \quad (2)$$

Where τ_{sim} represents the threshold for d distance in which blocks are considered similar. Further, both FW_{sim} and $FW_{dr}^{s,t}$ represents the search window in the term $I_{s,t}$ of size $n_d \times n_d$ centered on the position of B . We then map and combine all similar blocks to perform a higher quality image.

4. HYBRID LIGHT FIELD SUPER-RESOLUTION

To get to the final super-resolution image, we take the merit of our recent approach LFHYSR, which is used directly after we apply LFBMSR. The method is based on a combination of two classical approaches I_{MISR} and I_{HYSISR} . For more details see [2] Formally, the LFBMSR equation is formed as:

$$E_{SR}(\mathbf{x}) = \sum_k \|\mathbf{y}^{(k)} - I_{MISR} + I_{HYSISR}\| \quad (3)$$

Where E_{SR} represents the energy function and $\mathbf{y}^{(k)}$ represents the VPs, the two approaches I_{MISR} and I_{HYSISR} represents the observation method and the hybrid example-based method. We generate the Vps $\mathbf{y}^{(k)}$ by factor 4x4, using directional adapting interpolation to get $\mathbf{x}^{(0)}$, and is up sampled up to the size of \mathbf{x} by using our LFHYSR by factor 8x8. The advantage of using this combination depends on the fact that none of the recent approaches have combined MISR and HYSISR and exploits them to LF images. The gain is important for LF cameras to enhance the resolution.

Algorithm: Light field Block Matching Super-resolution

Input: Viewpoint images (VPs) as $\mathbf{y}^{(k)}$ generated from the plenoptic RAW data

1. To enhance the VPs resolution, do the following
 - a) Apply directional adapting interpolation
1. To apply LFBMSR do the following
 - b) Extract the depth information from the $\mathbf{y}^{(k)}$
 - c) Construction of blocks $n_a \times n_a \times k \times k$
 - d) In the reference VPs, we search for the set of blocks similar to B
 - e) Aggregation or mapping and 4D transform
2. To process LFHYSR do the following
 - f) Apply (MISR) I_{MISR} for method one
 - g) Apply (HYSISR) I_{HYSISR} for method two
3. To blend the two methods, the following is done
 - h) Pixel values of both methods are added to I_{HYSISR}
4. To segment and process the object's mask I_{HYSISR}
 - i) For each pixel in I_{HYSISR} identify sharp pixels I_{score}
 - j) From pixels in I_{score} generate clusters
 - k) Clusters are all connected to one contiguous region I_{app}
 - l) All neighbors of similar color create set of regions R_{color} in each pixel in I_{app}
 - m) Relevance score is updated for each region in R_{color}
 - n) To generate segmented image the noise is eliminated
5. To super resolve the segmented image apply SRR

Output: Super Resolution Image \mathbf{x}

5. EXPERIMENTAL RESULTS

In our experiment, we compared our LFBMSR performance's approach with the recent outstanding methods LFSR approaches and the evaluation is done taking into account the objective quality and visual impression. We perform our experiment from real data (*i.e.*, raw images), $L(u, v, s, t)$ two of which were obtained by our plenoptic camera, **Numbers** and **Toy**, whereas the **Knights** image was taken by Stanford University, which are reconstructed with size 9×9 VPs. All VPs resolution were set to 2248×1488 . For testing we have used several views from the raw image (*i.e.*, 3x3, 4x3, 5x4 and 5x5) as it is a free choice between the 9x9 views extracted. Each VP is partitioned into 4D macro blocks, where the maximum displacement within the search space is ± 7 pixels in horizontal as well as vertical directions for all data used (Numbers, Toy, Knights).

During implementation, we had an error while increasing the window size FW , (*i.e.*, 11×11 and 23×23); and we tested the pixels' behavior to find the best block match in all views to determine if the problem was in block matching or depth information. By using general block matching and the method of BMSR, we developed binary images and exploited them to identify the pixels to see which pixel defines the best match for each view. Then, we applied that again without constraint. This was repeated for all competed methods, MISR, HYSISR and the method in [14]. Because the depth can change when moving right or left, some pixels are not visible because of occlusion; this can confuse the block matching. Thus, the depth estimation method we used is too random. Therefore, we utilized a different depth estimation method from [9], which is better and meaningful. Thus, the window size is set to ± 7 pixels. The outcome results have improved after fixing the error. The quality of our method is measured by creating a ground truth image from LR images by the VPs such that each VP has 281×186 image resolution. The Matlab (R2014b) environment is utilized for all the experiments. The quality is evaluated with competing approaches in terms of PSNR, SSIM. The method was compared to our recent LFHYSR approach [2], bicubic interpolation, Wanner and Goldluecke [14], MISR [12] and HYSISR [13]. For visual comparison results, see Figure 3.

Table 1: Results of the average PSNR and SSIM (5x5 Views)

Methods	PSNR(dB)			SSIM		
	Toy	Numb.	Knights	Toy	Numb.	Knights
Bicubic	34.27	34.71	34.90	0.9211	0.9354	0.9367
HYSISR	35.53	35.73	35.86	0.9417	0.9388	0.9396
MISR	35.70	35.96	36.21	0.9429	0.9413	0.9423
Wanner <i>et al</i> [14]	36.29	36.44	36.67	0.9434	0.9438	0.9456
LFHYSR	36.45	36.71	36.82	0.9437	0.9498	0.9510
LFBMSR (Ours)	36.73	36.93	37.11	0.9458	0.9521	0.9532
Gain over MISR	1.03	0.97	0.90	0.0029	0.0108	0.0109

6. CONCLUSION

Our paper has implemented a novel LFBMSR method. The results were evaluated to the recent LFSR algorithms and compared with respect to objective quality as well as visual impression and outperform all comparison state-of-the-art methods with a total average achieved over MISR with respect to PSNR in tune of up to 1.06 dB and 0.97 dB; as well as 0.0108 and 0.0029 for SSIM, respectively, for the 'Numbers' and 'Toy' images, and 0.90 and 0.0109 for 'Knights' image. This is achieved by searching for the best similar blocks in all VPs and super resolving the final image using our LFHYSR approach.

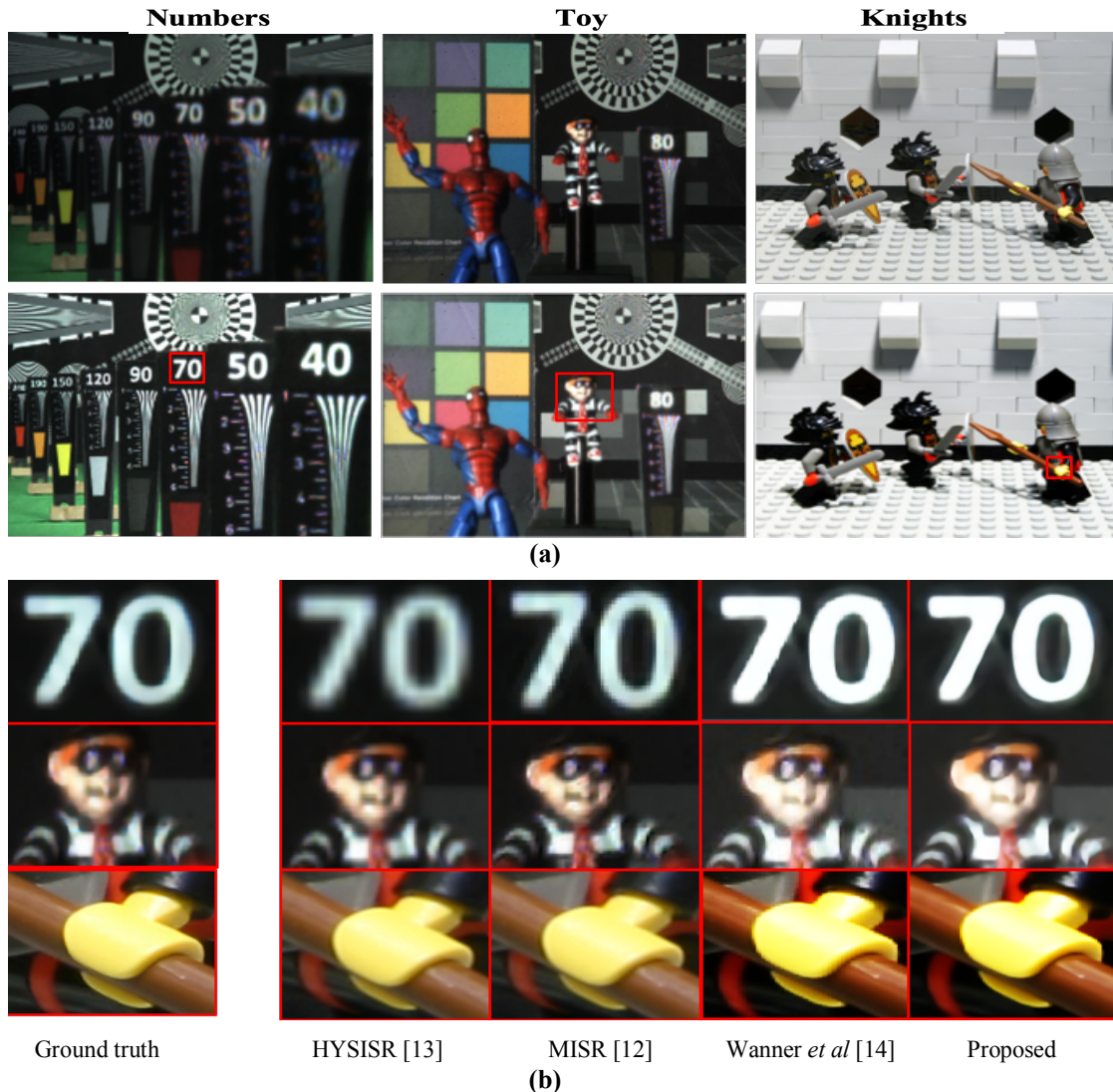


Figure 3. (a) Shows central view point images extracted from RAW images (Upper). Enhanced resolution image obtained after using our proposed LFBMSR method (Lower). (b) A close-up of all competitive super-resolved images.

7. REFERENCES

- [1] V. Velisavljevic, "Image interpolation using oriented wavelet transform", 16th *EUSIPCO*, August 2008.
- [2] S. Farag, V. Velisavljevic, and A. Aggoun, "A Hybrid Approach for Image Super-Resolution of Light Field Images," *IEEE Int. MMSP* October 2017.
- [3] M. Levoy. and P. Hanrahan. "Light Field Rendering. *ACM Trans. Graph.* (SIG- GRAPH), pages 31–42. 1996.
- [4] R. Ng, "Digital light field photography," *Stanford Univ.* 2006.
- [5] L. Peng and D. Liu, "All-in-focus image reconstruction based on plenoptic cameras," *Proc. - 2013 7th Int. ICIG 2013*, 2013.
- [6] E. H. Adelson and J. Y. A. Wang, "Single Lens Stereo with a Plenoptic Camera," *IEEE Trans. Pattern Anal. Mach. Intell.* 1992.
- [7] T. G. Georgiev and A. Lumsdaine, "Superresolution with Plenoptic 2.0 Cameras," *Front. Opt. 2009/Laser Sci. XXV/Fall 2009 OSA Opt. Photonics Tech. Dig.*, no. 2, p. STuA6, 2009.
- [8] T. E. Bishop, S. Zanetti, and P. Favaro, "Light field superresolution," *IEEE Int. ICCP 09*, 2009.
- [9] N. Sabater, M. Seifi, V. Drazic, G. Sandri, and P. Pérez, "Accurate disparity estimation for plenoptic images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8926, pp. 548–560, 2015.
- [10] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, 2012.
- [11] M. Alain, A. Smolic "Light Field Denoising by Sparse 5D Transform Domain Collaborative Filtering, V-SENSE project" *IEEE Int. MMSP* October 2017.
- [12] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, 2003.
- [13] Y. Xian, X. Yang, and Y. Tian, "Hybrid Example-based Single Image Super-Resolution," pp. 1–12. 2013
- [14] S. Wanner and B. Goldluecke, "Variational light field analysis for disparity estimation and super-resolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 3, pp. 606–619, 2014.
- [15] Y. Wang, G. Hou, Z. Sun, Z. Wang, and T. Tan, "A simple and robust super resolution method for light field images," *2016 IEEE Int. Conf. Image Process.*, pp. 1459–1463, 2016.
- [16] Y. X. Licheng Yu, H. Xu, and X. Yang, "Self-Example Based Super-Resolution With Fractal-Based Gradient Enhancement" Licheng Yu , Yi Xu , Hongteng Xu , Xiaokang Yang Shanghai Key Lab of Digital Media Processing and Communication." pp. 456–461, 2016.
- [17] E. Marenzi, E. Torti, F. Leporati, E. Quevedo, and G. M. Callicò, "Block matching super-resolution parallel GPU implementation for computational imaging," *IEEE Trans. Consum. Electron.*, vol. 63, no. 4, pp. 368–376, 2017.
- [18] K. Nasrollahi and T. B. Moeslund, *Super-resolution: A comprehensive survey*, vol. 25, no. 6. 2014.
- [19] A. Danielyan, R. Foi, V. Katkovnik, and K. Egiazarian, "Image and video super-resolution via spatially adaptive blockmatching filtering," *Proc. Int. Work. Local Non-Local Approx. Image Process.*, no. July 2014, p. 8, 2008.
- [20] Z. Li, H. Baker, and R. Bajcsy, "Joint image denoising using light-field data," *Electron. Proc. 2013 IEEE Int. Conf. Multimed. Expo Work. ICMEW 2013*, 2013.
- [21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.