A SPATIO-ANGULAR FILTER FOR HIGH QUALITY SPARSE LIGHT FIELD REFOCUSING

Martin Alain, Aljosa Smolic

V-SENSE Project, School of Computer Science and Statistics, Trinity College, Dublin

ABSTRACT

The ability to render synthetic depth-of-field effects post capture is a flagship application of light field imaging. However, it is known that many existing light field refocusing methods suffer from severe artefacts when applied to sparse light fields, known as angular aliasing. We propose in this paper a method for high quality sparse light field refocusing based on insights from depth-based bokeh rendering techniques. We first provide an in-depth analysis of the geometry of the defocus blur in light field refocusing by analogy with the defocus geometry in a traditional camera using the thin lens model. Based on this analysis, we propose a filter for removing angular aliasing artefacts in light field refocusing, which consists in modifying the well known shift-and-sum algorithm to apply a depth-dependent blur to the light field in between the shift and the sum operations. We show that our method can achieve significant quality improvements compared to existing approaches for a reasonable computational cost.

***Index Terms*—** Light field imaging, refocusing, angular aliasing, bokeh

1. INTRODUCTION

Light field imaging allows to capture all light rays passing through a given amount of the 3D space [1, 2], especially capturing angular information which is lost in traditional 2D imaging systems. We focus in this paper on the common two-plane parameterisation of light fields, in which the light field can be represented as a 4D function: $\Omega \times \Pi \rightarrow \mathbb{R}, (s,t,u,v) \rightarrow p(s,t,u,v)$, where the plane Ω represents the spatial distribution of light rays, also called the image plane, indexed by (u,v), while Π, the camera plane, corresponds to their angular distribution indexed by (s,t). In practice, the light field parameterised with two-parallel planes consists in a regularly sampled 2D grid of 2D images. The regular grid spacing on the camera plane is called the baseline, denoted b, while the 2D images are named sub-aperture images (SAI). We consider in this paper the variables s,t,u,v to be metric, and define their corresponding scalar indices i,j,k,l, where i,j are camera indices and k,l are pixel indices. For convenience, we define $L(i,j,k,l) \triangleq p(s,t,u,v)$ and we denote the SAIs by $I_{i,j}(k,l) \triangleq L(i,j,k,l)$.

Applications of light fields notably include rendering novel images viewpoints [1, 3], estimating scene geometry in the form of disparity or depth maps [4–6], and synthetic depth-of-field rendering or refocusing [7, 8]. In this paper, we focus on the latter application, for which many methods have been proposed. The shift-and-sum algorithm [7, 9] is a simple and a well known method to produce refocused images from a light field, in which the light field SAIs are first shifted towards the target focal plane and then averaged. An extension of this concept to the Fourier domain was later proposed in [10]. More advanced filters in the 4D Fourier domain have then been proposed to perform volumetric refocusing [8]. More recently, the Fourier Disparity Layer representation has been proposed [11], which allows rendering and refocusing in real time by exploiting parallelisation capabilities of modern GPUs.

However, the light field refocusing methods cited above exhibit artefacts when applied to sparse light field inputs, known as angular aliasing. The formal definition of densely (and by opposition sparsely) sampled light fields is given in the study of plenoptic sampling [12–14]. In [12], Chai et al. first provided guidelines for dense light field sampling. By considering the disparity between neighbouring SAIs, the condition for having a densely sampled light field is that its disparity should not exceed 1. Such condition is difficult to respect in practice, and many existing light field datasets are not strictly dense, in particular when captured with a gantry or a camera array. Therefore, multiple approaches have been developed to address angular aliasing in light field refocusing. A direct approach consists in reconstructing a dense light field from the sparse input before refocusing [15]. In order to avoid having to reconstruct a full dense light field or perform any pre-processing of the light field, Xiao et al. [16] proposed a method to detect angular aliasing using a statistical analysis of the refocused light field, and reduce the aliasing by using lower resolution versions of the refocused image from a Gaussian pyramid, which are then fused with Poisson image editing techniques [17]. Wang at al. proposed to use depth-based bokeh rendering methods (discussed below) in order to avoid angular aliasing artefacts, which is also combined with super-resolution of the in-focus region to render the final image [18]. A learning based method was recently proposed in which the angular aliasing filtering is considered as a denoising problem solved with a convolutional neural network [19].

Before light field refocusing, rendering synthetic bokeh has been a long standing application in computer graphics [20–22]. By analysing the geometry of the defocus blur in traditional cameras using the thin lens model, the radius of the circle of confusion (CoC), can be expressed depending on the aperture radius, the depth of the light point source, the depth of the focal plane, and the lens focal length. Given an all-in-focus input image, its corresponding depth map and the camera parameters, a synthetic

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 780470.

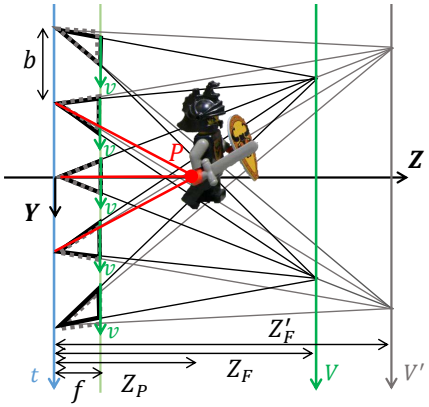

Fig. 1: The light field two-parallel plane parameterisation, explicitly defining the *local* and *global* parameterisation of the image plane. Re-parameterising the global focal plane from Z_F to Z'_F induces a shift of the local image planes, such that the relation of Eq. 1 always holds.

bokeh image can be rendered by applying a depth-dependant blur on the input image [20], where the blurring kernel is computed from the CoC radius. Note that such method has mainly been applied to synthetic images for which corresponding ground truth depth maps are known in order to simulate shallow depth-of-field, and real-time solutions have been developed e.g. for application in video games [23]. A well known method to speed up synthetic depth-of-field image rendering is to use a blurred lower resolution image from a Gaussian pyramid, called mipmap, for large CoC radius [21]. This technique is more difficult to apply to natural images, for which accurate depth maps are difficult to obtain, and is in fact used as an application to demonstrate the quality of estimated depth map in computer vision research [24], or requires manual editing [25].

We propose in this paper to combine insights from depth-based bokeh rendering methods with light field refocusing, in particular to address angular aliasing artefacts for refocusing of sparse light fields. We first provide in section 2 an in-depth analysis of the geometry of the defocus blur in light field refocusing by analogy with the defocus geometry in traditional cameras based on the thin lens model. Thanks to this analysis we can derive the angular aliasing conditions for light field refocusing based on the light field disparity, which confirms results given in the study of plenoptic sampling. Based on the previous study, and given that accurate disparity maps can be estimated from light fields, we propose in section 3 a novel refocusing method which can be seen as an extension of depth-based bokeh rendering methods applied to light fields. More precisely, the depth-based blurring is distributed in between the light field SAIs, which allows to reduce angular aliasing artefacts. We show in section 4 that the proposed approach can drastically improve the rendered image quality compared to classical and more recent approaches when applied to sparse light fields, for a reasonable computational cost.

2. TWO-PARALLEL PLANE LIGHT FIELD COC

We introduce in this section our analysis of the CoC of the refocused two-parallel plane light field. For this purpose we first need to define the parameterisation of the two-parallel plane light field in detail.

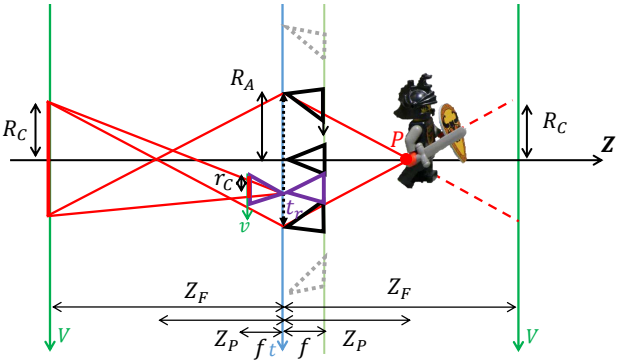

Fig. 2: Analysis of the CoC in light field refocusing. Note that the global image plane is mirrored with respect to the camera plane for clarity and by analogy with the defocus geometry in the thin lens model. The refocus camera positioned at (s_r, t_r) is displayed in purple.

2.1. Global and local image plane parameterisation

In this paper we make a distinction between a *global* focal plane, which is generally considered the image plane in common two-parallel plane parameterisation, and the *local* image planes, which correspond to the local image coordinates of a light field camera. Using this distinction, the position of the global focal plane is defined as the distance where the light field camera frustums coincide (see Fig. 1). We refer to this distance as the focal distance denoted Z_F. We denote the global focal plane coordinates as U, V and the local image coordinates as u, v, and we have the following relationship for all light field cameras:

$$u = Uf/Z_F,\ v = Vf/Z_F \tag{1}$$

where f is the focal length of the light field cameras. Note that re-parameterising the global focal plane, i.e. changing the focal distance, induces a shift of the local image planes with respect to the optical axis, as shown in Fig. 1, such that the relation in Eq. 1 always holds. This is unlike most definitions of the two-parallel plane light field in existing papers, which implicitly assume that all principal points are aligned with the camera optical axis (which corresponds to a focal plane at infinity in our case).

In such configuration, the relationship between the depth of a point P, denoted Z_P, and the corresponding disparity d_P between two neighbouring cameras can be expressed as:

$$Z_P = \frac{bf_{px}}{d_P + bf_{px}/Z_F} \tag{2}$$

where b is the light field baseline, and f_{px} is the focal length in pixel, i.e. $f_{px} = f/\varepsilon$ where ε is the pixel size in meter.

2.2. Light field circle of confusion

In this section we analyse the defocus blur in light field refocusing, and provide a precise formula for the CoC radius based on

known light field parameters. As shown in Fig. 2, we study the defocus geometry by analogy with the thin lens model, using the parameterisation described in the previous section. The radius of the CoC on the global focal plane R_C can be easily obtained depending on the aperture radius R_A using similar triangles:

$$R_C = R_A \frac{|Z_F - Z_P|}{Z_P} \tag{3}$$

The CoC radius on the local image plane of the refocus camera r_C can then be obtained thanks to Eq. 1:

$$r_C = R_A \frac{|Z_F - Z_P|}{Z_P} \frac{f}{Z_F} \tag{4}$$

We emphasise that this result holds for any position of the target refocus camera (s_r, t_r) on the camera plane, thanks to the precise parameterisation described in the previous section, as illustrated on Fig. 2 where the refocus camera is shown in purple. By dividing by the pixel size we get the CoC radius in pixel unit ρ_C:

$$\rho_C = \frac{R_A f_{px}}{Z_F} \frac{|Z_F - Z_P|}{Z_P} \tag{5}$$

which is similar (yet not equal) to the expression used in depth-based bokeh rendering methods [20,21]. By substituting Eq. 2 in the previous equation, we can express the CoC depending on the disparity instead of depth, which reduces to:

$$\rho_C = \rho_A |d_P - d_F| \tag{6}$$

with $\rho_A \triangleq R_A/b$. Advantageously, this equation does not depend on any of the metric parameters of the light field (focal length, baseline, focal plane distance) but only the disparity, which can be estimated from the light field SAIs directly. It is also easier in practice to use the scalar aperture radius ρ_A which corresponds to a number of cameras within the aperture, rather than the metric radius R_A.

From this analysis we can also observe that aliasing artefacts will occur in the defocus bokeh if the CoC radius ρ_C becomes larger than the aperture radius ρ_A, as there will not be enough light field cameras in the aperture to cover the area within the CoC (see Fig 3). Using Eq. 6 we can express this condition in terms of the disparity only:

$$|d_P - d_F| > 1 \tag{7}$$

As mentioned in the introduction, such condition is already known from previous work on plenoptic sampling [12–14]. We can see from Eq. 7 that to minimize the aliasing, the ideal placement for the focal plane is $d_F = (d_{max} - d_{min})/2$, which is also consistent with the results from the plenoptic sampling theory. However, contrary to the plenoptic sampling theory which is usually concerned with image-base rendering for all-in-focus images, by definition of the refocusing application d_F is not fixed, and thus prone to more severe angular aliasing.

In addition, we can also derive from Eq. 6 a condition on the disparity corresponding to the in-focus region of the refocused image, which occurs if the CoC radius is less than half a pixel, which we can express as:

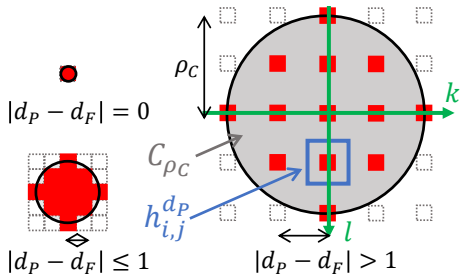

Fig. 3: Geometry of the light field CoC on the image plane and angular aliasing conditions. Top left: ray is in focus. Bottom left: Defocus with natural bokeh. Right: defocus with angular aliasing. The gray area in C_{ρ_C} correspond to parts of the bokeh not covered by any light field ray. We compensate for the missing rays by blurring known rays (in red) using kernels $h_{i,j}^{d_P}$.

$$|d_P - d_F| \leq \frac{1}{2\rho_A} \tag{8}$$

3. PROPOSED METHOD

Assuming that the light field is Lambertian, the core idea behind the proposed anti-aliasing filter is to compensate for missing rays in the aliased defocus bokeh, by blurring the known light field rays such that they fill the gaps in the target aperture shape (represented in gray in Fig. 3). At a high level, our approach consists in modifying the shift-and-sum algorithm by applying a disparity dependent blur to each SAI after the *shift* but before the *sum* operation. While pre-filtering the light field has been proposed in previous work, especially in the Fourier domain [8], our proposed spatio-angular filter is the first that directly relates to the shape of the circle of confusion.

As for the shift-and-sum, we first start by shifting all the SAIs in the light field toward the target refocus disparity d_F and the target refocus position i_r, j_r:

$$I_{i,j}^{d_F}(k,l) = I_{i,j}\big(k + (i - i_r)d_F,\, l + (j - j_r)d_F\big),\ \forall i,j \tag{9}$$

We then define for each light field camera position a disparity dependent blur kernel $h_{i,j}^{d}$ as:

$$h_{i,j}^{d}(k,l) = \begin{cases} C_{\rho_C}(k,l)/|C_{\rho_C}| & \text{if } |k-i| \leq |d - d_F| \\ & \&\ |l-j| \leq |d - d_F| \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where d is the disparity value, C_{ρ_C} is the disk inscribed in the CoC of radius ρ_C (computed using Eq. 6), and $|C_{\rho_C}|$ is the area of C_{ρ_C} used for normalisation. This normalisation factor simulates the dispersion of light on the image sensor in the thin lens defocus model.

The blurring is then applied on regions of the shifted SAIs corresponding to disparity value lying outside of the in-focus disparity range defined in Eq. 8:

$$I^n = \sum_{d=d_{start}}^{d_{end}} \sum_{i,j} (I_{i,j}^{d_F} \times \alpha_d) * h_{i,j}^d$$
$$I^d = \sum_{d=d_{start}}^{d_{end}} \sum_{i,j} \alpha_d * h_{i,j}^d \quad (11)$$

where I^n and I^d are numerator and denominator respectively, used to avoid any pixel intensity overflow, α_d is a binary mask identifying pixels corresponding to the disparity value d, $\times$ represent the pixel-wise multiplication, and $*$ represents the convolution operator. We apply the blurring independently to the background and the foreground, for which the disparity ranges are defined as $d_{start} = d_{min}$, $d_{end} = -1/(2\rho_A)$ and $d_{start} = 1/(2\rho_A)$, $d_{end} = d_{max}$ respectively. We thus obtain the numerator and denominator I_b^n, I_b^d for the background and I_f^n, I_f^d for the foreground.

We then generate an all-in-focus image at the target refocus position I_{i_r,j_r}. Note that if the position (i_r, j_r) does not correspond to an existing light field camera position, I_{i_r,j_r} can be obtained with a view synthesis method, e.g. using disparity warping. We also generate the corresponding disparity map D_{i_r,j_r}.

The final refocused image I_r is obtained by compositing the blurred background, the all-in-focus image, and the blurred foreground, following compositing rules from back to front:

$$I_r = \frac{(I_b^n(1-\alpha_F) + I_{i_r,j_r}\alpha_F)(1-\alpha_f) + I_f^n\alpha_f}{(I_b^d(1-\alpha_F)+1)(1-\alpha_f) + I_f^d\alpha_f} \quad (12)$$

where α_F is a mask identifying pixels corresponding to the in-focus region, i.e. $\alpha_F = |D_{i_r,j_r} - d_F| \leq 1/(2\rho_A)$, and α_f is mask identifying pixels corresponding to the blurred foreground, i.e. $\alpha_f = I_f^d > 0$. A more detailed summary of the proposed algorithm is shown in Algorithm. 1.

Following fast depth-based bokeh rendering approaches [21], we further propose to use a multi-scale approach to reduce complexity, such that the size of the blurring kernels $k_{i,j}^d$ is constant. For each shifted SAI $I_{i,j}^{d_F}$, a pyramid with n levels is created, with a downscaling factor σ. The number of levels for the pyramid can be determined depending on the maximum disparity range as $n = \log_\sigma(\max(|d_{min} - d_F|, |d_{max} - d_F|)$. During the disparity dependent blurring described in Eq. 11, for large blurring kernel size, i.e. large d, the blurring can be applied on higher level of the pyramid, which can thus be performed with a smaller kernel. For a given disparity d, the adequate pyramid level l can be computed as $l = \log_\sigma(d)$, for which the disparity is reduced to $d\sigma^l$ and we can use the corresponding blurring kernel $h_{i,j}^{d\sigma^l}$.

Note that some existing methods fully rely on such a multi-scale approach to reduce the angular aliasing, i.e. use blurred images from higher pyramid levels instead of applying a spatial blur as proposed in this paper. However, the specific spatial blurring kernels we propose are essentials to faithfully simulate the target aperture shape, which can be lost when directly using the blurred pyramid images, especially for light fields with low angular resolution. Furthermore, while we represent circular aperture in our figures for simplicity, our method can be used for any target aperture shape.

Algorithm 1 Disparity-based light field bokeh rendering

procedure LFREFOCUS(L, d_F, ρ_A, i_r, j_r)
  Generate all-in-focus image at target position I_{i_r,j_r}
  Generate corresponding disparity map D_{i_r,j_r}
  $d_{min} \leftarrow \min(D_{i_r,j_r})$, $d_{max} \leftarrow \max(D_{i_r,j_r})$
  Generate shifted light field L^{d_F} using Eq. 9 ▷ Shift
  $I_b^n, I_b^d \leftarrow$ DISPBLUR($L^{d_F}, D_{i_r,j_r}, d_{min}, \frac{-1}{2\rho_A}, \rho_A$)
  $I_f^n, I_f^d \leftarrow$ DISPBLUR($L^{d_F}, D_{i_r,j_r}, \frac{1}{2\rho_A}, d_{max}, \rho_A$)
  $\alpha_F \leftarrow |D_{i_r,j_r} - d_F| \leq 1/(2\rho_A)$
  $\alpha_f \leftarrow I_f^d > 0$
  Compute I_{i_r,j_r}^r using Eq. 12 and return.
procedure DISPBLUR($L, D, d_{start}, d_{end}, d_F, \rho_A$)
  $I^n \leftarrow 0$, $I^d \leftarrow 0$
  $d \leftarrow d_{start} - d_F$
  for $d = d_{start}$ to d_{end} **do**
    $\rho_C \leftarrow \rho_A|d - d_F|$
    Compute $h_{i,j}^d$ using Eq. 10
    $\alpha_d \leftarrow d \leq |D - d_F| \leq d + 1$
    $B_{i,j} \leftarrow (I_{i,j} \times \alpha_d) * h_{i,j}$ ▷ Blur
    $\alpha_{i,j} \leftarrow \alpha_d * h_{i,j}^d$
    $I^n \mathrel{+}= B_{i,j}$ ▷ Sum
    $I^n \mathrel{+}= \alpha_{i,j}$
  Return I^n, I^d

4. EXPERIMENTS AND RESULTS

In this section, we compare our proposed approach, with the multi-scale implementation (denoted as MS) and without, to the shift-and-sum (SAS) algorithm [7], the Fourier Disparity Layer (FDL) [11], the selective light field refocusing (SLFR) approach of [18], and our multi-scale method applied to a single image, which can be seen as a single image depth-based bokeh rendering (SIBR) similar to [21]. For the FDL and the SLFR we use the Matlab implementation provided by the authors with the recommended parameters. Note that the FDL is using the GPU capabilities of Matlab. We implemented all the other methods ourselves in C++, with parallelisation on CPU. In all experiments, our multi-scale approach is used with a downscaling factor $\sigma = 0.5$ to create the SAI pyramids. The estimated disparity maps used in our experiments were obtained with an approximate but fast approach described in [5], for which is a C++ implementation is provided by the authors.

4.1. Objective evaluation

In order to perform an objective evaluation, we generated reference refocused images using the shift-and-sum algorithm applied on the full light field, and test sparse light fields are obtained by subsampling the original light fields. The refocused images are rendered for the center view, with a circular aperture with radius ρ_A equal to half the camera array size, and the target focus disparity $d_F = (d_{max} - d_{min})/2$. In this experiment we use 11 light fields from the *additional* subset of the synthetic HCI benchmark dataset [26] which have a resolution of $9 \times 9 \times 512 \times 512$, and 9 light fields from the Stanford gantry dataset [27] which have a resolution of $17 \times 17 \times 1024 \times 1024$ (the spatial resolution can vary). We use subsampling factors α_s of 2, 3 and up to 4 for

Table 1: Average scores for the HCI dataset. Best and second best scores are highlighted in **bold** and *italic* respectively. GT indicate the use of the ground truth disparity maps.

Method	PSNR / SSIM	
	$\alpha_s = 2$	$\alpha_s = 4$
SAS [7]	**47.23 / 0.997**	36.14 / 0.971
FDL [11]	40.12 / 0.977	33.81 / 0.940
SLFR [18] - GT	x	36.14 / 0.985
SIBR [21] - GT	36.85 / 0.984	34.08 / 0.973
Ours - GT	*42.64 / 0.995*	**39.84 / 0.991**
Ours w/ MS - GT	41.85 / 0.994	*37.93 / 0.988*
SLFR	x	35.03 / 0.975
SIBR	34.80 / 0.974	33.92 / 0.970
Ours	39.39 / 0.987	37.68 / 0.983
Ours w/ MS	39.06 / 0.986	36.76 / 0.983

the Stanford dataset. Note that by design the selective refocusing approach of [18] only uses 3×3 SAIs from the input light fields, which corresponds to the maximum subsampling factor. Average PSNR and SSIM scrores are given for the two datasets in Tables 1 and 2.

As ground truth disparity maps are available for the HCI dataset, we provide in Table 1 results obtained with the ground truth disparity (denoted GT) as well as estimated disparity for comparison. While there is a clear decrease in performance for all disparity dependant methods when using the estimated disparity maps compared to the ground truth, our proposed approach is still outperforming all other methods for the sparsest input setting ($\alpha_S = 4$). For $\alpha_S = 2$, our method using ground truth disparity map is second best after the shift-and-sum algorithm, and second best in terms of SSIM without ground truth disparity.

The experiment on the Stanford dataset confirms that our proposed approach outperforms existing methods for sparse input light fields, as shown in Table 2.

We can observe for both datasets that our multi-scale implementation suffers from a decrease in performance compared to the single scale implementation, however we can see from the average processing times reported in Table 3 that it can be advantageous in terms of computational complexity, especially for the Stanford dataset which has a higher resolution. We can also observe from Table 3 that while the proposed approach has an increased complexity compared to the shift-and-sum algorithm, it is faster than the single image depth-based bokeh rendering. The FDL being computed on GPU is clearly the fastest method, and we postulate that our approach could reach similar processing time thanks to a GPU implementation.

4.2. Visual results

We apply the different test methods to light fields from the Technicolor dataset [28], captured with a camera array with a wide baseline. In addition, the spatial resolution is also much higher than for the test light fields used in the previous section, and thus the disparity range of these light fields is around 50 pixels. This is much larger than the disparity range of previous test light fields, even subsampled, and thus more challenging.

We compare in this experiment our proposed multi-scale approach to the shift-and-sum algorithm, the Fourier Disparity Layer, and the selective light field refocusing approach. We show results for the "Painter" light field in Fig. 4. While some artefacts can be observed in the refocused images obtained with our method due to errors in the disparity maps, our approach clearly provides better results than existing methods.

Table 2: Average scores for the Stanford dataset. Best and second best scores are highlighted in **bold** and *italic*.

Method	PSNR / SSIM		
	$\alpha_s = 2$	$\alpha_s = 4$	$\alpha_s = 8$
SAS [7]	**42.68 / 0.988**	*36.13* / 0.954	28.85 / 0.893
FDL [11]	*42.40* /0.986	35.86 / 0.950	28.66 / 0.881
SLFR [18]	x	x	29.36 / 0.931
SIBR [21]	31.64 / 0.964	31.48 / 0.963	29.38 / 0.951
Ours	39.59 / *0.987*	**38.18 / 0.982**	**32.97 / 0.964**
Ours w/ MS	39.31 / 0.986	35.01 / *0.977*	*31.20 / 0.961*

Table 3: Processing time in seconds.

Method	HCI		Stanford		
	$\alpha_s = 2$	$\alpha_s = 4$	$\alpha_s = 2$	$\alpha_s = 4$	$\alpha_s = 8$
SAS [7]	0.08	0.05	0.80	0.34	0.20
FDL [11]	0.01	0.01	0.03	0.02	0.02
SLFR [18]	x	10.79	x	x	65.35
SIBR [21]	0.46	0.31	5.66	3.74	3.22
Ours	0.53	0.79	17.33	17.72	22.15
Ours w/ MS	0.36	0.21	3.99	1.58	0.85

Fig. 4: Visual results for a sparse light field captured with a camera array [28] for the different test methods. The approximate average processing time is shown next to each method.

5. CONCLUSION

We proposed in this paper a novel refocusing method for sparse light fields based on the in-depth analysis of circle of confusion for the two-parallel plane parameterisation and borrowing insight from depth-based bokeh rendering methods. We showed through our results that our method can outperform classic refocusing algorithms as well as more advanced methods proposed in recent years.

However, our results also showed that the quality of our re-

sults depends on the quality of the disparity maps estimated from the input light field. In future work, we are planning to investigate variations of the proposed method robust to disparity estimation errors, e.g. using more advanced alpha matting techniques to blend the blurred background, foreground and all-in-focus image.

In addition, the proposed approach can also be used for more applications than just angular aliasing filtering, e.g. simulate small apertures radius ($\rho_A < 1$) for sparse light fields, or very large aperture for dense light fields. Our method could also be extended to re-parameterised light fields for tilt-shift refocusing.

6. REFERENCES

[1] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. ACM SIGGRAPH*, 1996, pp. 31–42.

[2] Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F Cohen, "The lumigraph," in *Proc. ACM SIGGRAPH*, 1996, pp. 43–54.

[3] Aaron Isaksen, Leonard McMillan, and Steven J. Gortler, "Dynamically reparameterized light fields," in *Proc. ACM SIGGRAPH*, 2000, pp. 297–306.

[4] Bastian Goldluecke, "Globally consistent depth labeling of 4D light fields," in *Proc. IEEE CVPR*, 2012, pp. 41–48.

[5] Y. Chen, M. Alain, and A. Smolic, "Fast and accurate optical flow based depth map estimation from light fields," in *Proc. IMVIP*, 2017.

[6] Yu-Ju Tsai, Yu-Lun Liu, Ming Ouhyoung, and Yung-Yu Chuang, "Attention-based view selection networks for light-field disparity estimation," in *Proc. AAAI Conference on Artificial Intelligence*, 2020, vol. 34, pp. 12095–12103.

[7] V. Vaish, B. Wilburn, N. Joshi, and M. Levoy, "Using plane + parallax for calibrating dense camera arrays," in *Proc. IEEE CVPR*, 2004, vol. 1, pp. I–I.

[8] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams, "Linear volumetric focus for light field cameras.," *ACM Trans. on Graphics*, vol. 34, no. 2, pp. 15–1, 2015.

[9] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light Field Photography with a Hand-Held Plenoptic Camera," Tech. Rep., Stanford University CSTR, Apr. 2005.

[10] Ren Ng, "Fourier slice photography," in *Proc. ACM SIGGRAPH*, 2005, pp. 735–744.

[11] Mikael Le Pendu, Christine Guillemot, and Aljosa Smolic, "A fourier disparity layer representation for light fields," *IEEE Trans. on Image Processing*, vol. 28, no. 11, pp. 5740–5753, 2019.

[12] Jin-Xiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum, "Plenoptic sampling," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 307–318.

[13] Cha Zhang and Tsuhan Chen, "Spectral analysis for sampling image-based rendering data," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 11, pp. 1038–1050, 2003.

[14] Zhouchen Lin and Heung-Yeung Shum, "A geometric analysis of light field rendering," *International Journal of Computer Vision*, vol. 58, no. 2, pp. 121–138, 2004.

[15] Chao-Tsung Huang, Jui Chin, Hong-Hui Chen, Yu-Wen Wang, and Liang-Gee Chen, "Fast realistic refocusing for sparse light fields," in *IEEE Proc. ICASSP*. IEEE, 2015, pp. 1176–1180.

[16] Zhaolin Xiao, Qing Wang, Guoqing Zhou, and Jingyi Yu, "Aliasing detection and reduction in plenoptic imaging," in *Proc. IEEE CVPR*, 2014, pp. 3326–3333.

[17] Patrick Pérez, Michel Gangnet, and Andrew Blake, "Poisson image editing," in *Proc. ACM SIGGRAPH*, pp. 313–318. 2003.

[18] Yingqian Wang, Jungang Yang, Yulan Guo, Chao Xiao, and Wei An, "Selective light field refocusing for camera arrays using bokeh rendering and superresolution," *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 204–208, 2018.

[19] Shachar Ben Dayan, David Mendlovic, and Raja Giryes, "Deep sparse light field refocusing," in *Proc. BMVC*, 2020.

[20] Michael Potmesil and Indranil Chakravarty, "A lens and aperture camera model for synthetic image generation," *Proc. ACM SIGGRAPH*, vol. 15, no. 3, pp. 297–305, 1981.

[21] Sungkil Lee, Gerard Jounghyun Kim, and Seungmoon Choi, "Real-time depth-of-field rendering using anisotropically filtered mipmap interpolation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 3, pp. 453–464, 2009.

[22] Neal Wadhwa, Rahul Garg, David E Jacobs, Bryan E Feldman, Nori Kanazawa, Robert Carroll, Yair Movshovitz-Attias, Jonathan T Barron, Yael Pritch, and Marc Levoy, "Synthetic depth-of-field with a single-camera mobile phone," *ACM Trans. on Graphics*, vol. 37, no. 4, pp. 1–13, 2018.

[23] Kleber Garcia, "Circular separable convolution depth of field," in *ACM SIGGRAPH Talks*, pp. 1–2. 2017.

[24] Jonathan T Barron, Andrew Adams, YiChang Shih, and Carlos Hernández, "Fast bilateral-space stereo for synthetic defocus," in *Proc. IEEE CVPR*, 2015, pp. 4466–4474.

[25] Dongwei Liu, Radu Nicolescu, and Reinhard Klette, "Stereo-based bokeh effects for photography," *Machine Vision and Applications*, vol. 27, no. 8, pp. 1325–1337, 2016.

[26] Katrin Honauer, Ole Johannsen, Daniel Kondermann, and Bastian Goldluecke, "A dataset and evaluation methodology for depth estimation on 4D light fields," in *Asian Conference on Computer Vision*. Springer, 2016.

[27] "The stanford light field archive," `http://lightfield.stanford.edu/lfs.html`, accessed: 11-03-2021.

[28] N. Sabater, G. Boisson, B. Vandame, P. Kerbiriou, F. Babon, M. Hog, R. Gendrot, T. Langlois, O. Bureller, A. Schubert, and V. Allié, "Dataset and pipeline for multi-view light-field video," in *Proc. IEEE CVPRW*, 2017, pp. 1743–1753.