



HAL
open science

Axial refocusing precision model with light fields

Zhaolin Xiao, Jinglei Shi, Xiaoran Jiang, Christine Guillemot

► **To cite this version:**

Zhaolin Xiao, Jinglei Shi, Xiaoran Jiang, Christine Guillemot. Axial refocusing precision model with light fields. *Signal Processing: Image Communication*, 2022, pp.1-11. 10.1016/j.image.2022.116721 . hal-03651953

HAL Id: hal-03651953

<https://hal.science/hal-03651953>

Submitted on 26 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Axial refocusing precision model with light fields

Zhaolin Xiao*, Jinglei Shi†, Xiaoran Jiang†, Christine Guillemot† (IEEE Fellow)

Abstract—In light field imaging, axial refocusing precision corresponds to the minimum distance in the axial direction between two distinguishable refocusing planes. The refocusing precision can be essential for applications like light field microscopy. In this paper, we introduce a refocusing precision model based on a geometrical analysis of the flow of rays within the virtual camera. The model establishes the relationship between the feature separability of refocusing and different camera settings. As extending numerical aperture (NA) in classical imaging, the baseline extension of light field also gives more accurate refocusing results. To test the axial refocus precision, we conduct experiments with 1st generation Lytro camera as well as a Blender light field simulation. The results is basically consistent with our prediction. Then, we show that computationally extending the light field baseline increases the axial refocusing precision on real plenoptic camera and light field microscopy datasets.

Index Terms—Light field, Computational photography, Axial resolution, Sampling analysis, Digital refocusing.

I. INTRODUCTION

LIGHT field imaging has become popular in the last years, due to its potential for a variety of applications. In particular, post-capture digital refocusing can be easily implemented by shifting and adding the sub-aperture images [1], [2], [3], [4], [5]. In theory, the light field focal plane can be moved continuously along the axial direction. However, in real systems, due to limited spatial resolution and baseline, the number of distinguishable focal planes is limited in the axial direction. In order to distinguish from the concept of NA in classical imaging, we use the baseline to denote the distance between two farthest separated angular views throughout this paper. In traditional imaging, the notion of depth of field (DoF) is a widely accepted metric for axial resolution, and for which a theoretical analysis can be found in [6]. The definition of DoF concerns the acceptable sharpness within a given axial range, but we are instead interested in the focus precision, *i.e.* the feature separability between each focal plane and its adjacent planes when applying digital refocusing.

In this paper, we introduce a light field axial refocusing precision (LF-ARP) model within a general virtual camera. This model allows us to calculate the minimum spacing between two distinguishable adjacent refocusing positions in the axial dimension, given the fixed focal length, the baseline of the main lens (or aperture size), the sensor size and its resolution, and the object plane distance. Please note that the considered refocusing precision differs from the notion of DoF, it does not only consider the in-focus regions, but also the out-of-focus blur. The out-of-focus blur is function of the distance between the objects and the current focal plane, which can

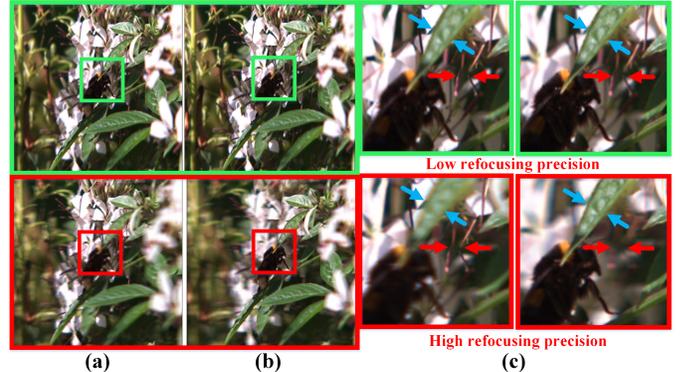


Fig. 1. Light field refocusing precision illustration. (a) a low refocusing precision case. A given disparity corresponds to a specific focal plane when applying digital refocusing, *b* represents the baseline in the angular plane. Using a 1.0X baseline, the difference between refocusing at $disparity = 0.0$ and $disparity = +0.5$ can be hardly perceivable. (b) High precision refocusing with the same focus setting as in (a), from which we can see clear differences with a computational 2.0X enlarged baseline. (c) On the right zoomed views, stamens and leaves can be clearly found lying on different depth layers with high refocusing precision.

affect the possibility to distinguish objects at different axial positions.

The refocusing precision can be essential for some applications, *e.g.* light field microscopy [7], [8], and light field particle image velocimetry (LF-PIV) [9], [10]. However, we show that the classical Fourier slicing refocusing method [11] or the spatial shifting-adding re-sampling technique [1] produce refocused images do not allow differentiating features within the predicted precision, as shown in Figures 1(a) and (b) where we used the same focal plane setting. When using the shift-and-add refocusing method on the original light field, one can hardly see refocusing differences in Figure 1-(a). On the contrary, Figure 1-(b) shows that, by enhancing the refocusing precision, *e.g.* by increasing the baseline, we can better distinguish the objects at different axial positions. Please notice that, while the light field baseline often denotes the spacing between two adjacent views (or cameras), here the term *baseline* will refer to the spacing between two farthest views (or cameras). To extend the angular baseline, we introduce a learning-based view extrapolation algorithm in [12]. As shown in Figure 1, if the different objects need to be distinguished via refocusing, one can optimize the capture configuration parameters, such as the focal length, the aperture size or the baseline, the distance to the object plane. Here, we present a geometrical analysis of the axial precision model, which reveals the relations between the axial refocusing precision and the camera parameters. Quantitative and qualitative experiments are also conducted using both real and synthetic light field datasets.

* Xi'an University of Technology, Xi'an 710048, China.

† Institut National de Recherche en Informatique et en Automatique (INRIA), Rennes 35000, France.

II. RELATED WORK

A light field can be seen as a sampling of the *plenoptic function* introduced by Adelson and Bergen in [13] to describe the radiance of light rays emitted by a scene along any direction, at any time and for every wavelength. By assuming that the light rays are not attenuated along a line, light fields are commonly represented as 4D functions with spatial and angular coordinates [14], [15].

Let $LF(x, y, u, v)$ be the 4D representation of a light field, in which \mathbf{XY} and \mathbf{UV} stand for the spatial and angular planes respectively. For sake of simplicity, in the rest of the manuscript, we consider a 2D slice $LF(x, u)$ of the light field. At capture time, physically changing the object distance or the focal length are two most straight forward operations to adjust image focus. Chang *et al.* [16] designed a focus-tunable lens with precise tracking of the focal length, which can scan 1600 focal planes per second.

Post-capture synthetic refocusing [1] can also produce a focal stack, i.e. a sequence of images each focused on a different plane, by re-sampling the input light field as

$$I_\alpha(x) = \int_u LF_0(x + \alpha u, u) \quad (1)$$

Therefore, one can generate different refocusing from the light field LF_0 by varying the parameter α , which implies the focal depth. Theoretically, one can set a focal plane Ω_0^α at a different axial position by tuning the parameter α . Figure 2(a) illustrates the effect of refocusing on a 2D slice of the light field, where X and U denote spatial axis and angular axis respectively. The disparity variation αu is 0 on the original focal plane. The shifting factor $\alpha \in [-\infty, +\infty]$, is positive or negative when we focus closer or further respectively from the original focus plane. This leads to different shearing angles in the sampling space, as shown in Figures 2(b) and 2(c).

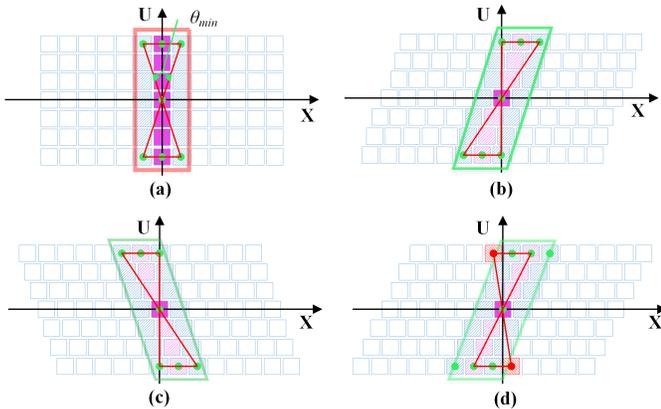


Fig. 2. Differences between original and re-sampled light fields. (a) The original 2D light field, (b) The positive shifting corresponds to a re-sampling of the original light field in (a), in order to move the focus plane closer in the axial direction. (c) The negative shifting is opposite to (b). (d) The shifting leads to a resolvable re-sampling even without sub-pixel interpolation.

However, the number of focal stack images, or slices in the resulting volume, is limited by the axial resolution of the light field. Indeed, if α is small, then even for the outmost view u_{max} , the disparity variation $\alpha * u_{max}$ may be less than 1 pixel.

In this case, without sub-pixel interpolation, the re-focused images obtained by the re-sampled light field will appear the same as the re-focused images computed from the original light field. Similarly, if the disparity of the outmost view u_{max} with respect to the central one is less than 1 pixel, then the shearing angle will be less than θ_{min} (red box in Figure 2-(a)), and the re-focused images with the original or re-sampled light fields will not show any difference without sub-pixel interpolation. If the shearing angle is larger than θ_{min} , it will be obviously resolvable because of the contribution of the new angular samples, as shown in Figure 2-(d) (marked in the slanted red line).

In classical imaging systems, increasing the numerical aperture (NA) decreases the DoF, but leads to a higher axial resolution (also called depth resolution) [17], [18]. Many solutions have been proposed to deal with the narrow DoF problem, such as image deblurring based on 3D PSF modeling and all-in-focus image fusion from multiple axial scans [19]. In the object space, Chen *et al.* [20] and Hahne *et al.* [21] proposed different optical models to accurately measure the distance of the object plane based on a geometric analysis of the standard plenoptic camera. Further, Hahne *et al.* [21] derived the distance to the refocused object plane and its corresponding DoF for different light fields, which has been experimentally verified by placing objects at the predicted distances.

Unlike the DoF distance in [21], the separability of LF-ARP refers to whether light field re-sampling will affect the integral imaging result, considering both the sharp focused regions and the blurred defocused regions. For each axial position, we derive the minimum light field re-sampling so that shifts of 1 pixel of the most extreme views, when constructing the focal stack with the shift-and-add procedure, yields distinguishable refocusing planes. The refocusing precision Arp is equivalent to the DoF only if the DoF of all the focal slices are non-overlapping, which is not verified in most of cases.

III. AXIAL REFOCUSING PRECISION

We start by defining the refocusing precision in light field imaging and by recalling the notion of DoF. Then, we describe the proposed light field axial refocusing precision (LF-ARP) model.

A. Refocusing precision: definition

Let Ω_0 be a given focal plane on the object side, $\Omega_0^{\alpha+}$ and $\Omega_0^{\alpha-}$ are the nearest distinguishable planes around Ω_0 from the far side and the near side respectively. The refocusing precision is defined as the minimum distance between two distinguishable adjacent focal planes $\Omega_0^{\alpha+}$ and $\Omega_0^{\alpha-}$ in the object space, and can be expressed as

$$Arp(LF_0, \Omega_0) = \left[d(\Omega_0^{\alpha-}), d(\Omega_0^{\alpha+}) \right] \quad (2)$$

s.t. $\|LF_0^\alpha - LF_0\|_\infty < \varepsilon$, if $\alpha \in [\alpha^-, \alpha^+]$

where $d(\Omega_0)$ stands for the distance between the target focal plane of Ω_0 (on the object side) and the camera plane (or the aperture plane) UV . Here, the UV plane is assumed to lay at

the zero axial position. The parameter ε denotes a negligible difference between the original light field LF_0 and the re-sampled one LF_0^α . Figure 3 shows the difference between the axial refocusing precision and the DoF in the object space, and shows that the two adjacent distinguishable focus planes and the image planes located at distances $d(\Omega_0^{\alpha^-})$ and $d(\Omega_0^{\alpha^+})$ from the camera plane do not correspond to the borders of the DoF. The DoF is the distance (red and blue dash lines), within which the object is in-focus, i.e. without optical blur. The DoF has been exaggerated compared to the dimension of the imaging system for better visibility. Due to finite and discrete sampling of the 4D light field, the

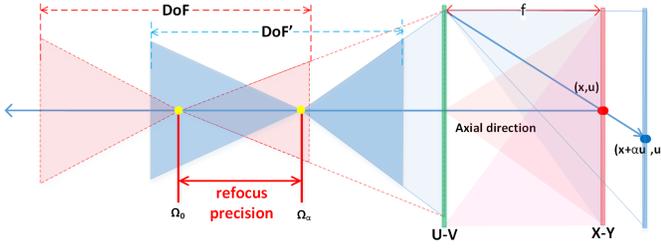


Fig. 3. Difference between light field DoF and refocusing precision. The pink and blue shadowed regions correspond to the definition of traditional DoF. The refocusing precision is defined as the minimum distance Arp between two distinguishable adjacent focal planes in the object space.

$Arp(LF_0, \Omega_0)$ cannot be infinitely small with a fixed spatio-angular resolution. Therefore, deriving the range $[\alpha^-, \alpha^+]$ for each possible focal plane Ω_0 is the main goal of this paper.

B. Depth of Field (DoF)

The axial re-focusing precision has some relationship with the concept of DoF. In traditional imaging, the DoF limit is given by

$$DoF = \frac{\lambda n}{NA^2} + \frac{n}{MNA} e \quad (3)$$

where λ stands for the light wave length, n represents the refraction index of medium, and NA is the numerical aperture of the entire imaging system. The variable e is the minimum resolvable distance in the image plane of the main lens, whose lateral magnification from object to image plane is M . Theoretically, the total DoF is determined primarily by wave optics (first term of Equation (3)) if the pixels are small enough to not limit resolution. But, in case of lower numerical apertures, the DoF is dominated by the geometrical optical circle of confusion (CoC) represented by the second term of Equation (3).

Please note that, if assuming the DoFs of all the focal slices are non-overlapping in a focal stack, only the in-focus regions are concerned, then the refocusing precision Arp is equivalent to the DoF . However, this assumption usually does not hold in real light field imaging systems for the two following reasons. First, since the pixel size may not be small enough to satisfy the theoretical wave DoF limit, the actual DoF of the different slices of the focal stack may be larger and overlapping in most practical systems, as shown in Figure 3. Second, the refocused

pixel size is $N_u \times N_v$ times larger than the sensor pixel size or the diffraction-limited spot size. Here, $N_u \times N_v$ is the angular resolution. Therefore, in the case of real light field imaging systems, and assuming $N_u = N_v$, the DoF of refocused image is given by [6]

$$DoF \approx \frac{(2 + N_u) \lambda n}{2NA^2} \quad (4)$$

where the geometrical optical CoC, which is $N_u/2$ times larger than the wave optics term, dominates. As a consequence, the DoF of light field imaging is significantly larger than the DoF of traditional imaging with the same resolution sensor. Therefore, in order to obtain a better spatial or lateral resolution, Broxton *et al.* [22] suggest to apply 3D-deconvolution on the defocused angular samples. The DoF resolution can also be enhanced by applying image based super-resolution, e.g. the state-of-art learning based light field super-resolution [23].

While the DoF only measures the range of in-focus area, our motivation here is to provide a model that measures the distance between adjacent distinguishable focus planes. Unlike the discussion on the distance of refocused object plane and the corresponding DoF in [21], the proposed model reveals the refocusing or re-sampling precision given a constant image plane and fixed field of view (FoV).

C. Axial refocusing precision model

In this section, we propose a refocusing precision Arp model, in which both the image plane and the FoV are assumed to be fixed.

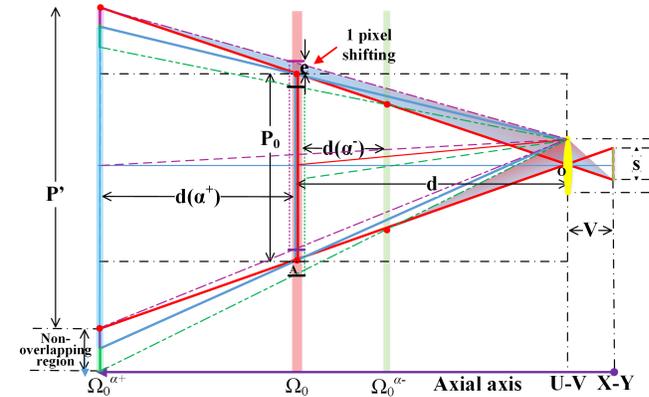


Fig. 4. Illustration of axial refocusing precision. The axial refocusing precision is a function of the object distance d , the focal length f , the angular baseline of the main lens b , and the sensor size s . Since a 1 pixel shift of the outmost view corresponds to the minimum possible re-sampling of the light field, for any possible object side focal plane Ω_0 , we can derive the distances to nearest resolvable adjacent refocusing positions $\Omega_0^{\alpha^+}$ and $\Omega_0^{\alpha^-}$ in both directions.

Assuming the image plane distance V to be constant when re-focusing with light field imaging, the following thin lens equation gives the relationship between the object distance d and the virtual focal length f after re-focusing

$$\frac{1}{d} + \frac{1}{V} = \frac{1}{f} \quad (5)$$

where the focal length f is a variable, and d denotes the distance between the virtual focal plane and the physical main lens.

Let us now consider re-focusing with the shift-and-add procedure. Let r_x and r_y represent the horizontal and vertical resolutions of the sensor, and b stands for the baseline width. Considering a 1 pixel shift, by geometrically analyzing Figure 4, one can deduce that the light field refocusing precision Arp , for any possible focal plane Ω_0 , can be expressed as

$$\begin{aligned} Arp(LF_0, \Omega_0) &= \phi(d, f, b, s) \\ &= [d - d(\alpha^-), d + d(\alpha^+)] \\ &= [d(1 - \frac{M'e'}{b/2 + M'e'}), d(1 + \frac{M'e'}{b/2 - M'e'})] \end{aligned} \quad (6)$$

where

$$e' = \frac{s(d-f)}{f\sqrt{r_x^2 + r_y^2}} \quad (7)$$

is the minimum distinguishable re-sampling distance in millimeters (mm), in the object side, corresponding to the diffraction limit $\frac{\lambda}{2NA}$ on the image side. In most of practical imaging systems, the pixel size is larger than this diffraction limit, so that the pixel size corresponds to the minimum re-sampling unit e on the focal plane, as shown in Figure 2(d). Figure 3 indeed shows that the 1 pixel shift of outmost views corresponds to the minimum resolvable unit distance e on the focal plane Ω_0 . As shown in Figure 4, $d(\alpha^-)$ and $d(\alpha^+)$ represent the distances between the backward and forward refocus planes and the focal plane, which are marked as the green, blue and red planes respectively. Please note that, $d(\alpha^-)$ and $d(\alpha^+)$ are different from $d(\Omega_0^{\alpha^-})$ and $d(\Omega_0^{\alpha^+})$, which are measured with respect to the coordinate origin of the optical center. The term M' is the reciprocal of lateral magnification factor M (i.e. from image side to object side) for the imaging system having a magnifying lens (e.g. a light field microscope). Please see more details of the model derivation in the Appendix VI.

The above LF-ARP model is a function of the object distance d , the focal length f , the width of baseline b , and the sensor size s . In traditional imaging, the sensor size s is constant, hence changing the focal length f implies changing the field of view (FoV). However, our model shows that, in the case of light field imaging, computational refocusing, i.e. changing the focus plane in which the object of interest is assumed to be, is equivalent to changing both the focal length f and the sensor size s , while keeping the physical distance between the image plane and the aperture plane constant, i.e. V is a constant. Figure 4 shows that a 1 pixel shift changes the convergence of the corresponding angular sample, as illustrated by the slight shifting orientations of the outmost view in purple, blue and green lines.

The concept of DoF can be seen as a metric based on focus and blur separation, in which the CoC varies with the shift of the image plane. Unlike the DoF, the proposed Arp model is derived assuming a given constant image plane V . More precisely, the FoV of P_0 and P' (i.e. the physical width of the virtual images on $\Omega_0^{\alpha^+}$ and Ω_0 respectively in the object space) stays the same when refocusing on different axial planes $\Omega_0^{\alpha^+}$ and Ω_0 , i.e. the image of the object, remains constant, when re-focusing.

Please note that it assumes a thin lens model and does not account for possible off-axis aberrations. All the variables of our LF-ARP model correspond to physical measures. If the parameters of equation (6) and equation (7) are known, one can easily compute the axial indistinguishable refocusing ranges for each object plane, for a given minimum possible re-sampling limit, e.g. 1 pixel size. The refocusing separability not only depends on the axial depth, but also on the scene texture. But, only the minimum refocusing separability distance is derived by considering the characteristic of camera part, the object scene texture frequency is assumed high enough to satisfy the resolvable requirements in this paper.

Equations (6) and (7) show that the Arp can be better with a close capture, or when using a lens with a long focal length. Figure 5-(a) shows how the distances $d(\alpha^-)$ and $d(\alpha^+)$ on the axial direction, and the Arp vary when re-focusing with a shift of the outmost view of multiple e units. We can indeed see that the variation of the distance $d(\alpha^-)$ contributes more to the Arp decrease than the one of $d(\alpha^+)$. We can indeed see that the blue curve is always above the green curve, which indicates that the distance $d(\alpha^+)$ contributes more to the Arp decrease than the one of $d(\alpha^-)$. This means that the physical distance between the current focal plane and the closest distinguishable adjacent focal plane in the axial direction will be larger when refocusing further, than when refocusing on a plane closer to the main lens.

Since the Arp decrease is non-linear as u increases, axial separability can be improved via spatial super-resolution [22], however the authors in [22] give depth-dependent band limits for spatial super-resolution, indicating that one can not recover the spatial resolution up to the band limit using sharp focused angular samples. Instead of applying spatial super-resolution, computationally expanding the baseline b can be another option to increase the Arp . As we show in Figure 5-(b), there is a correlation between the Arp and the baseline width b that following the equations (6).

IV. EXPERIMENTS

In this section, we test the refocus precision model considering synthetic scenes generated using the light field plugin of the Blender software [24]. The quantitative and qualitative comparisons are made using multiple synthetic light fields, and real light fields captured by 1st generation Lytro cameras, as well as on light field microscopy datasets.

A. Testing of the refocus precision model

To meet the resolvable requirements, ideally, we would need to capture a 3D scene with highly textured objects full filling the FoV, and then axially shift the object plane. Instead, for this first experiment, we created a synthetic experimental scene assuming a capture using a 9*9 camera array, with parameters ideally set as shown in Fig.6, i.e. focal length $f = 100.0mm$, baseline $b = 42.65mm$, multiply factor $M' = 1.0$, sensor size $s = 35.0mm$ and its resolution $r_x = r_y = 512$.

For a first validation of the proposed Arp model, we created a synthetic scene using Blender. The scene consists of two clipped standard resolution cards (RCs) placed at different

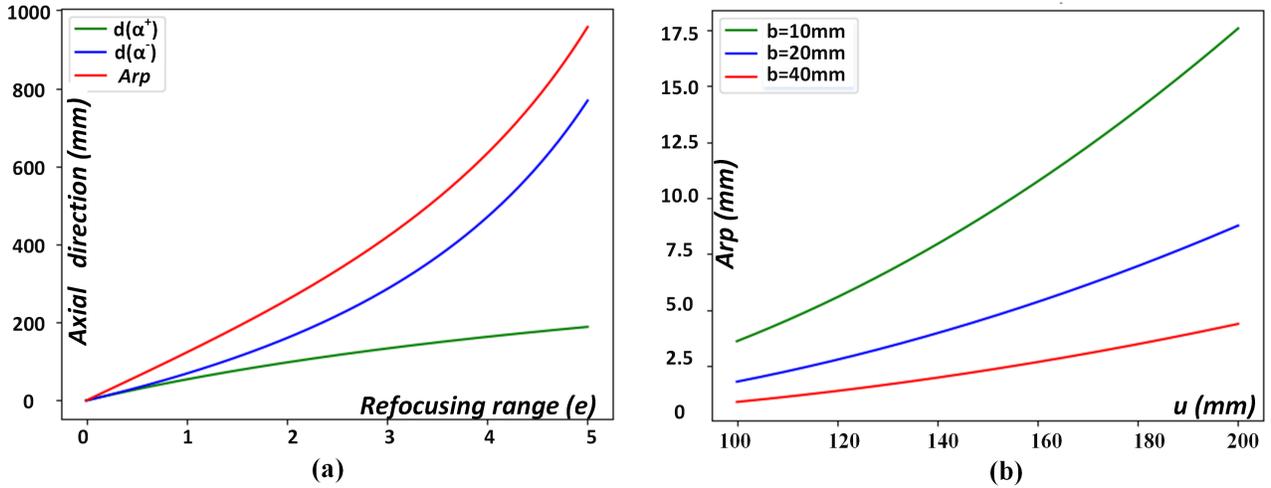


Fig. 5. LF-ARP for different parameter settings. (a) $d(\alpha^-)$, $d(\alpha^+)$ and Arp when re-focusing corresponding to a shift of multiple e units of the outmost view. Here, we set $f = 30\text{mm}$, $s = 28\text{mm}$, $M' = 1.0$, $r_x = r_y = 512$, $b = 10\text{mm}$, $d = 500\text{mm}$. The Arp range is the sum of the blue curve and the green one, which correspond to $d(\alpha^-)$ and $d(\alpha^+)$ respectively. (b) Arp as a function of the object distance d for different values of the baseline b . The setting of f , s , M' and r_x, r_y is the same as in (a).

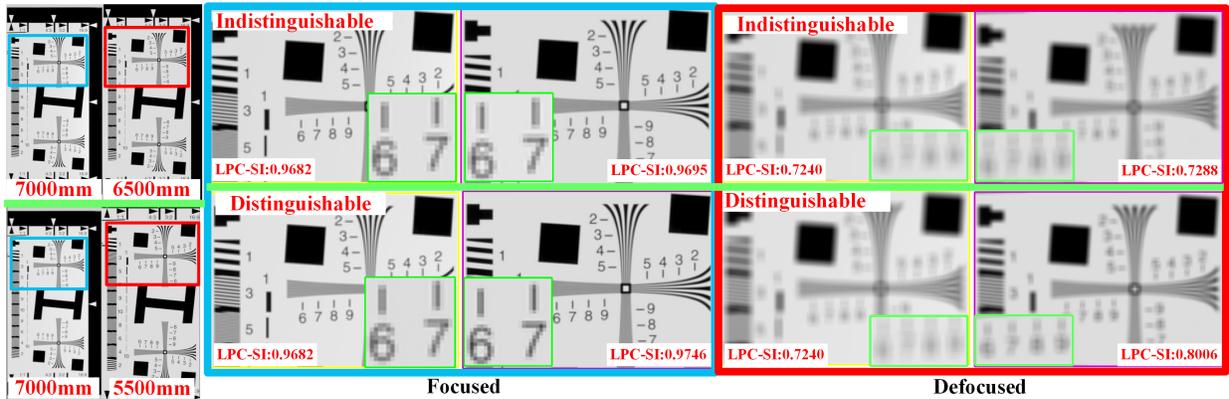


Fig. 6. Refocusing visual comparison when two boards are separated by a distance that is lower or higher than the LF- Arp value using Blender simulations. In the experiment, the estimated $d(\alpha^-)$ is equal to 6053 mm with a refocusing made at 7000 mm. For this reason, the resolution board is first placed at 6500 mm (within the Arp range, in the top row), and then at 5500 mm (out of the Arp range, in the bottom row). We show the zoomed refocusing results when the resolution boards are in focus, or out-of-focus respectively. From 8.0x zoomed views (the green boxes), we can see clearly the separability changing between the two different separations, especially for the defocused case.

positions in the axial direction (see the left part of Fig.6). We placed a reference RC at a distance of 5.0mm, 6.0mm and 7.0m from the UV plane (*i.e.*, the aperture plane) respectively, while the second RC is placed at neighbouring positions. The placement of the second RC at different neighboring positions allows us to measure refocus differences between focus planes in the same focal stack. Assuming a fixed image plane V , we calculated the ideal Arp range by varying only f , s .

For validating the model, we performed re-focusing at different planes, and then we compared the refocusing results of the regions corresponding to the two RCs, at five different axial positions, closer and further away from the reference focus plane, as shown in table I. We computed the focus/sharpness measure of the image regions of the two RCs. These measures should be very close when the re-focus at the two positions in the axial direction is not distinguishable. In order to measure

the separability of the two RCs, we employed the local phase coherence-based sharpness index [25] (LPC-SI). Instead of PSNR or SSIM, we use LPC-SI because it needs no reference image. Another reason is that the LPC-SI is less sensitive to the variation of scale and FoV as moving the second RC. In a same focal stack, let FS_{ref} be the image region of the reference RC, which is placed at the original object plane at d_0 , and FS_m be the image region of the compared movable RC, which can be placed at a neighbouring axial position d_k . As in equation 8, we consider a new quantitative measure of the focal stack based refocusing separability $RDiff(\cdot)$, which measures sharpness and blur of multiple focal slices as

$$RDiff(d_k) = \frac{\bar{SI}(FS_m(d_k), FS_{ref})}{\bar{SI}(FS_m(d_0), FS_{ref})}, \quad (8)$$

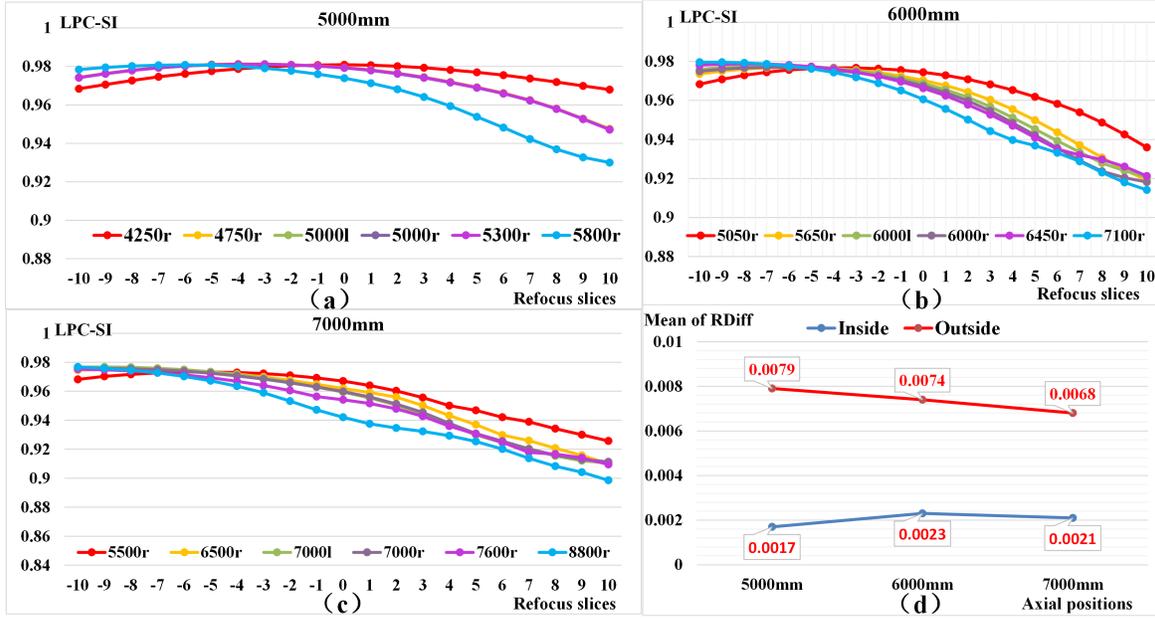


Fig. 7. Quantitative focus measurements when placing the resolution board inside and outside the Arp at different axial positions. (a)-(c) are focus measurements using LPC-SI. In each subplot, the red line and blue line show the distinguishable differences of refocusing when putting the resolution board outside the LF-Arp. The green lines are reference refocusing when setting the resolution board at 5000 mm, 6000 mm, and 7000 mm. If putting the resolution board inside the LF-Arp, the refocusing difference is hardly distinguishable, as shown by the other lines. (d) shows the mean difference between the reference refocusing and the refocusing inside or outside LF-Arp.

where

$$\bar{SI}(FS_a, FS_b) = \frac{1}{2N+1} \sum_{i=-N}^N (SI(FS_a^i) - SI(FS_b^i))^2 \quad (9)$$

and where $SI(\cdot)$ is the LPC-SI sharpness measure as defined in [25] (LPC-SI), FS_a^i and FS_b^i are two different image regions in the i -th focal slices respectively.

For each pair of the reference and compared RCs at $\langle d_0, d_k \rangle$, we first draw the LPC-SI curves when both RCs are placed at the original object plane d_0 , e.g. the 5000l (the reference one) curve and 5000r (the compared one) curve in Fig.7 (a). From Fig.7, we can see that the two curves are highly consistent when the two RCs are placed at the reference object plane. The consistency implies that the refocusing of the two RCs is indistinguishable for all focal slices, in both the sharp focused and blurred defocused regions. When placing the compared RCs away from the reference position and outside the Arp range, their refocusing becomes more distinguishable, especially when it is severely defocused (see the red and blue curves in Fig.7). When keeping the compared RCs inside the Arp, the curves are similar with the reference one. A visual comparison is shown in Fig.6.

The refocus precision is decreasing when the reference object plane moves away from the UV plane. We can indeed see that the curves become closer to each other, as shown in Fig.7 (b) and (c). Fig.7 (d) shows that the mean of $RDiff$ is 3 to 4 times higher, compared to the positions outside and inside the Arp range.

In order to further test the refocusing precision, we simulate a natural scene, which is composed of multiple pieces of complex textures. Different from the two RCs scene, as shown

in the top right sub-images of Fig. 8, the scene is assumed to be captured at the distance $d_0 = 2000mm$, with the focal length $f = 30.0mm$, baseline $b = 113.14mm$ and sensor size $s = 28.0mm$. These parameters have been ideally set according to the camera viewpoint.

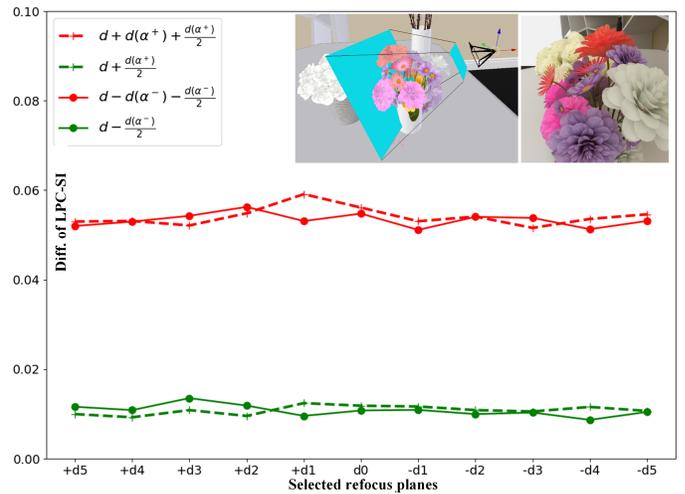


Fig. 8. Validation of the proposed refocus precision model. We show the LPC-SI between images refocused on adjacent planes within and outside the refocus precision limit $Arp = [d - d(\alpha^-), d + d(\alpha^+)]$ at different axial positions. On the top right, the synthetic scene with ideal camera settings is generated using the Blender light field plugin [24] with the parameters shown in Table II.

Fig. 8 shows that, within the refocus precision limit $Arp = [d - d(\alpha^-), d + d(\alpha^+)]$, with interval near and far boundaries denoted $d - d(\alpha^-)$ and $d + d(\alpha^+)$ respectively, the LPC-SI difference between two refocused images on adjacent planes

is very small (green curves), while it becomes much larger (red color lines) when the spacing between two refocus planes is larger than the Arp interval. The dashed and solid lines indicate the LPC-SI variation when increasing and decreasing the refocus distance respectively. From Fig. 8, we can see that the axial separability or resolvability is significantly higher when the refocus plane spacing is larger than the Arp limit, which indicates the correctness of the proposed Arp model. Please note that, there is no standard metric for the measurement of axial separability or resolvability, which not only depends on the characteristics of the camera, but also on the scene texture.

Fig.9 shows our experiment setup for testing the refocus precision of the 1st generation Lytro camera. The reference RC is placed in front of the camera, at distances of $500mm, 600mm, 700mm, 800mm, 900mm, 1000mm$ from the aperture plane respectively. The second RC is then placed at neighbouring positions using a translation device, which can be precisely controlled in steps of $\pm 0.01mm$ by a digital controller. The RCs are placed perpendicular to the optical axis of the camera, so that the axial distance between the camera aperture plane and the RCs can be simply measured using a laser rangefinder. In order to keep the illumination constant, we use a 200 W LED photographic light at a color temperature of 6500 K. The Arp limit will be less than $1mm$ at a distance of 200mm or even closer. Therefore, we choose to measure the Arp limit at a range from $500mm$ to $1000mm$. The movable RC is placed at reference positions from 10 to 20 times a step of $1mm$ on both sides. The setting of the Lytro camera is manually fixed while capturing a same reference position, *i.e.* the reference RC is constant. In real scene experiments, we employ the same refocusing precision metric, and set $RDiff(\cdot) = 2.5$. In the experiments, we use a focal length $f = 35.0mm$, a baseline $b = 17.5mm$ and sensor size $s = 7.67mm$, $r_x = 3280$, $r_y = 3936$, $M = 1.0$ as the ideal parameters of the 1st generation Lytro camera. Due to the non-ideal illumination and sensor noise, the captured real scenes are more blurred and noisy than the simulated ones. Therefore, we consider smaller clipped regions of both RCs for computing the LPC-SI. Fig. 10 (a)-(c) shows that the peak of the red curve (outside the Arp range) is obviously different from the other curves. Due to the noise and blur of the sampling stage, the LPC-SI values are within a smaller range than for the simulated scene, which leads to a lower separability in refocusing. The predicted and measured Arp values of the Lytro camera along the axial direction are shown in Fig. 10 (d). Due to the presence of noise and distortion, and the limited number of tested discrete axial positions, there is an obvious difference between the measured and predicted values of Arp , *e.g.* at the positions of $600mm, 800mm, 900mm$ and $1000mm$. Nevertheless, we can still see the high similarity in trend and values from the curves, which indicates the correctness of the proposed Arp model. The proposed Arp model is only validated with discrete re-focus planes positions. Validating the model for a continuum of positions on the optical axis would require designing a scene with carefully chosen distribution of objects, texture for each object, camera exposure settings, etc.

B. Evolution of refocusing precision with the baseline

By extending the baseline (or the aperture), the digital refocusing can produce a shallower DoF, thus leading to more accurate refocusing. To validate the axial refocusing enhancement, we further tested the resolvability or separability of the adjacent refocusing results. Different from the LPC-SI metric, the SSIM is employed to measure the differences between two adjacent refocus images. In Figure 11, for each axial focal plane (corresponding to different disparities), we give the mean SSIM of 7 Stanford microscopy datasets. A small value of mean SSIM indicates high separability. The black curve stands for separability for each axial focal plane with the ground truth 1.0X baseline light field, while the other curves marked with cross and triangles represent the axial separability of refocused results with 2.0X and 4.0X baselines respectively. We can see that the curves follow the same trend as those in Figure 5 (b), in the sense that the Arp value is lower when increasing the baseline, *i.e.* that enables a more precise refocusing. And the refocusing separability is decreasing in a non-linear manner when moving the focal plane further away from the original conjugate focal plane Ω_0 at axial position d_0 on the object side, as also shown in Figure 5 (a).

Fig. 11 shows the curves of the extrapolation method proposed in [26] based on multi-plane images (MPI) and of the shearing and extrapolation network (SENet) we proposed in [12], in green color and red color respectively, from which we can see that even the computational baseline extension can give better performances than the 1.0X baseline in the axial separability. The vertical peak in Figure 11 can be explained by the non-linear characteristics of the disparity space. When away from the zero disparity plane, the distance between the distinguishable adjacent refocus planes increases, for the same disparity value. A larger physical separation will lead to high separability (smaller SSIM value), which yields the central peak of SSIM curves. The left and right wings of the curves are due to the fact that the entire specimen is out of focus, the blurriness indeed leads to a high SSIM value. Finally, Figure 12 visually compares refocused images with 1.0X and 4.0X baseline. The detail changes can hardly be recognized with 1.0X baseline original refocusing, *e.g.* , those areas pointed by the green arrow. But, one can easily distinguish the sharp focused plane of these details on the refocused images with 4.0X baseline. To extend the baseline to more than 4.0X, we suggest interested readers to use other state-of-art learning based solutions, *e.g.* . the solutions in [27], [28].

C. Limitations

The proposed refocus precision analysis assumes a thin lens model, and does not account for possible off-axis deformations or distortions. The proposed Arp model is validated only at discrete and limited axial positions. Although the LPC-SI-based refocus sharpness metric does not need a known reference like PSNR or SSIM, it is not a standard metric by considering only the characteristics of the refocus difference. Therefore, a potential better metric will be still helpful, especially as in presence of noise.

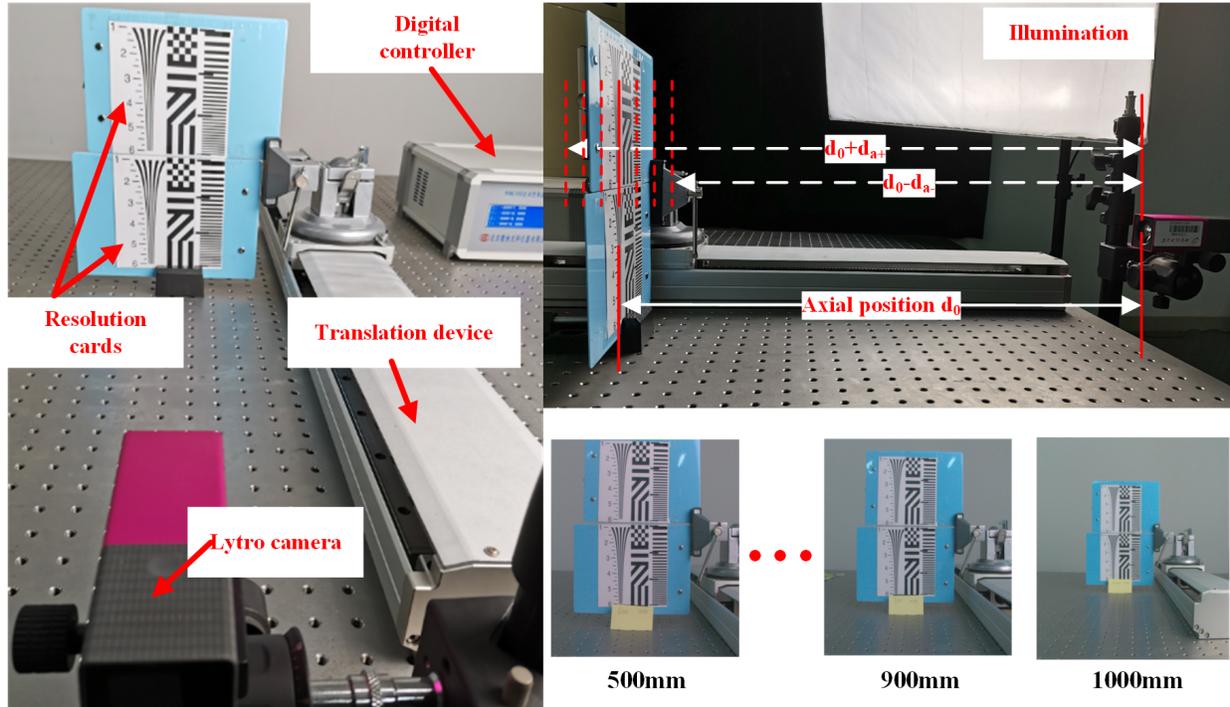


Fig. 9. Experimental setup of refocus precision evaluation for the 1st generation Lytro camera at different axial positions.

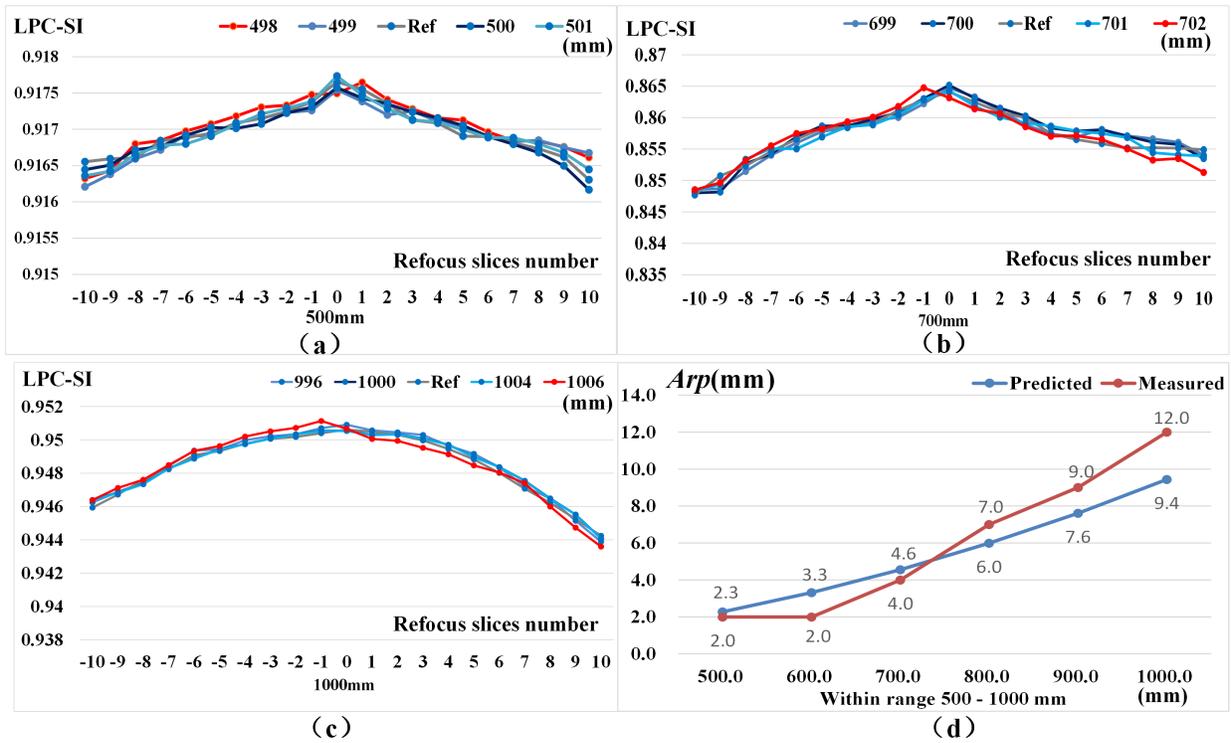


Fig. 10. Quantitative evaluations of refocus precision of Lytro camera along axial direction. (a)-(c) show the quantitative focus measurements when changing the spacing between the reference resolution board and the moving one. The red line corresponds to the case where refocusing is outside Arp, while the refocusing inside Arp corresponds to the remaining lines. (d) By testing Arp on discrete and limited positions, we show a comparison between ideal and measured Arp of Lytro camera within the range of 500mm-1000mm.

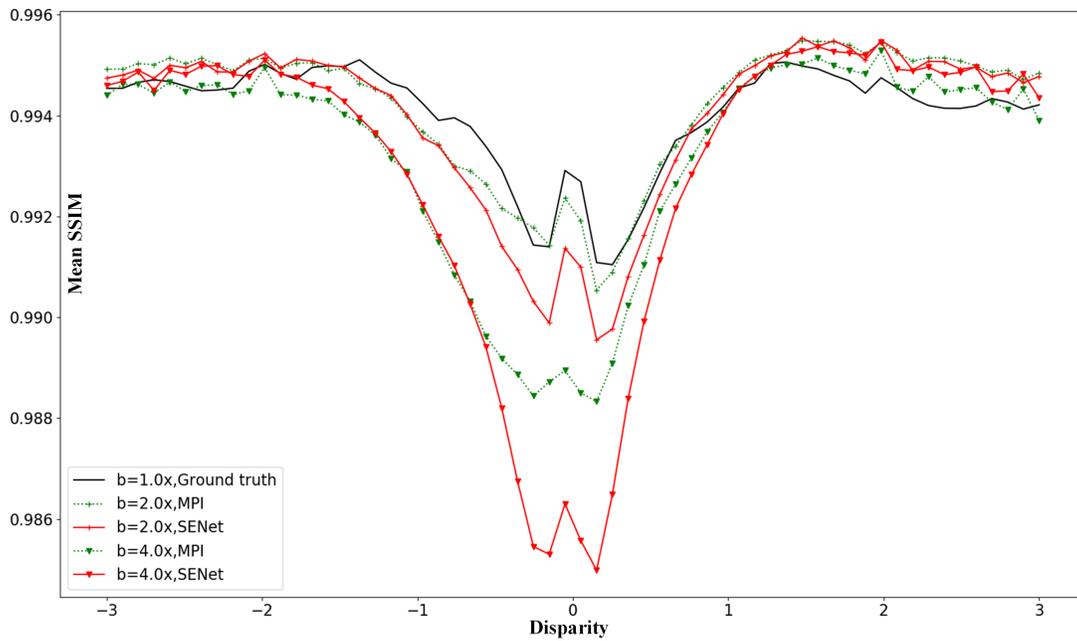


Fig. 11. SSIM between two adjacent refocus images with 1.0X, 2.0X, 4.0X baseline extension factors for two methods, the MPI-based method [26] and the proposed one. Lower is the SSIM, more distinguishable (i.e. showing more structural differences) are the two adjacent refocus planes.

V. CONCLUSION

In this paper, we introduced a light field refocusing precision model, which establishes the minimal resolvable or distinguishable spacing in the axial direction, given the fixed focal length, baseline, sensor size and resolution, and object plane distance. While the concept of DoF can be seen as a metric based on focus and blur separation, the proposed geometrical analysis is derived in the spatio-angular sampling space, and measures the distance between adjacent distinguishable focus planes. The model also reveals the relationship between the refocusing precision and camera settings. The proposed model can be used to optimize some configuration parameters of the capturing settings, such as the distance to the object plane, the aperture size or the baseline. It can benefit potential light field applications with specific axial refocusing precision demands, e.g. microscopy, particle velocimetry etc.

ACKNOWLEDGMENT

This work has been funded in part by the National Natural Science Foundation of China (No. 61871319 and No. 62031023), in part by the Natural Science Basic Research Plan of Shaanxi Province(No.2019JM-221), in part by the China Scholarship Council (CSC) under grant No.201808610055, and in part by the EU H2020 Research and Innovation Programme under grant agreement No 694122 (ERC advanced grant CLIM)

DISCLOSURES

The authors declare no conflicts of interest.

REFERENCES

- [1] A. Isaksen, L. McMillan, and S. J. Gortler, "Dynamically reparameterized light fields," in *ACM SIGGRAPH*, 2000, pp. 297–306.
- [2] G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, and Y. Liu, "Light field image processing: An overview," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 926–954, 2017.
- [3] I. Ihrke, J. Restrepo, and L. Mignard-Debise, "Principles of light field imaging: Briefly revisiting 25 years of research," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 59–69, 2016.
- [4] G. Wetzstein, I. Ihrke, D. Lanman, W. Heidrich, R. Raskar, and K. Akeley, "Computational plenoptic imaging," *Computer Graphics Forum*, vol. 30, no. 8, pp. 2397–2426, 2011.
- [5] M. Levoy, "Light fields and computational imaging," *Computer*, vol. 39, no. 8, pp. 46–55, 2006.
- [6] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," *ACM TOG*, vol. 25, no. 3, pp. 924–934, 2006.
- [7] R. Prevedel, Y. G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrödel, R. Raskar, M. Zimmer, and E. S. Boyden, "Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy," *Nature Methods*, vol. 11, no. 7, pp. 727–730, 2014.
- [8] L. Palmieri, G. Scrofani, N. Incardona, G. Saavedra, M. Martínez-Corral, and R. Koch, "Robust depth estimation for light field microscopy," *Sensors*, vol. 19, no. 3, pp. 1–16, 2019.
- [9] J. Belden, T. T. Truscott, M. C. Axiak, and A. H. Tchet, "Three-dimensional synthetic aperture particle image velocimetry," *Measurement Science and Technology*, vol. 21, no. 12, p. 125403, 2010.
- [10] C. Skupsch and C. Brücker, "Multiple-plane particle image velocimetry using a light-field camera," *Optics Express*, vol. 21, pp. 1726–1740, 2013.
- [11] R. Ng, "Fourier slice photography," in *ACM SIGGRAPH*, 2005, pp. 735–744.
- [12] Z. Xiao, J. Shi, X. Jiang, and C. Guillemot, "A learning-based view extrapolation method for axial super-resolution," *Neurocomputing*, vol. 455, pp. 229–241, 05 2021.
- [13] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, 1991, pp. 3–20.
- [14] M. Levoy and P. Hanrahan, "Light field rendering," in *ACM SIGGRAPH*, 1996.
- [15] S. Gortler, R. Grzeszczuk, R. Szeliski, and F. Cohen, "The lumigraph," in *ACM SIGGRAPH*, 1996.

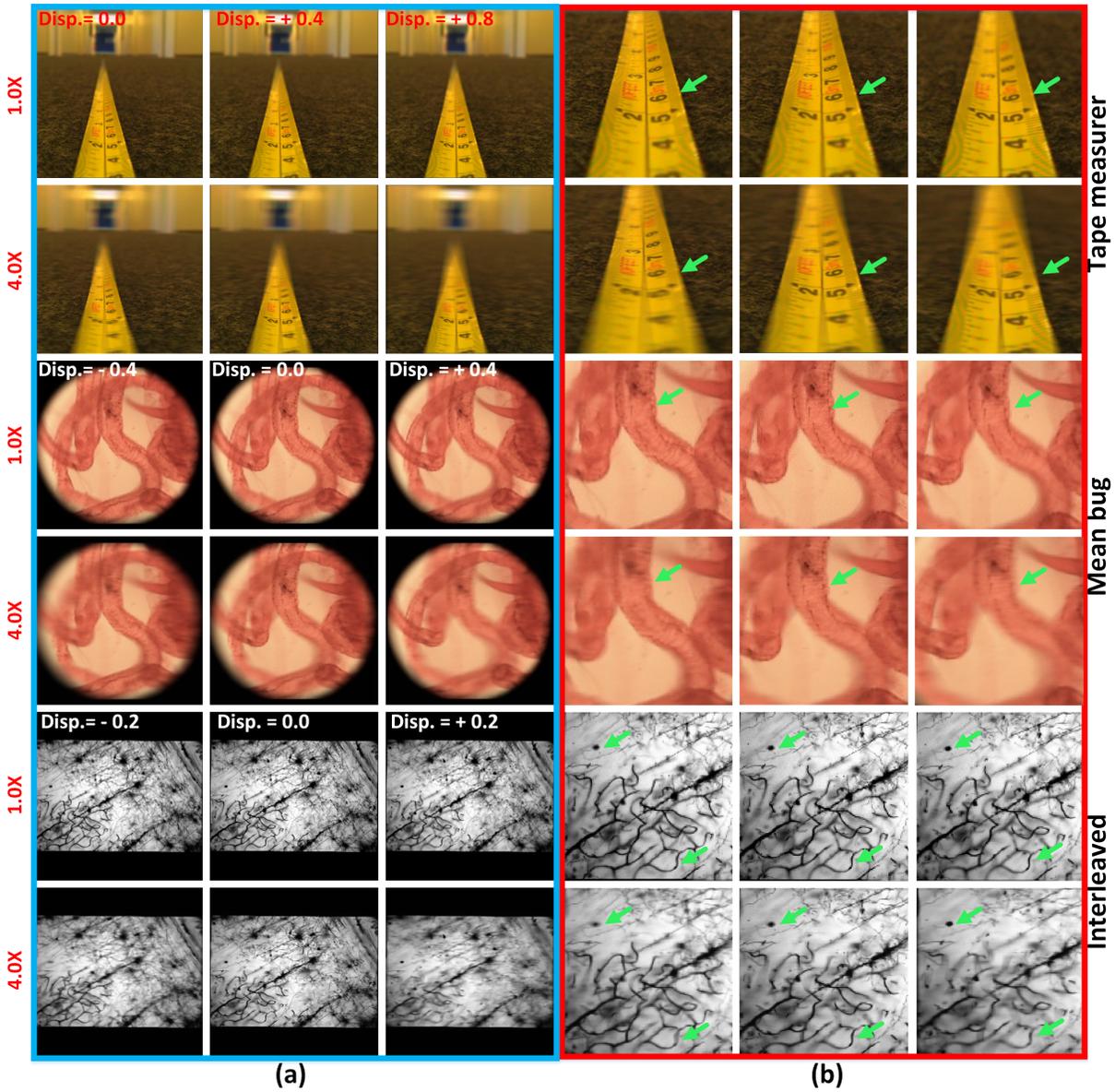


Fig. 12. Comparison of axial refocusing results with 1.0X, 4.0X baselines, where different disparity values correspond to different focal planes in axial direction. (a) Refocused results on same axial focus planes, with 1.0X and 4.0X baselines respectively. (b) Zoomed views, the green arrows are pointing at details on different refocusing planes that are more distinguishable with a 4.0X baseline and less resolvable with the 1.0X baseline.

TABLE I
PARAMETERS OF THE IDEAL LIGHT FIELD CAMERA ARRAY FOR TESTING LF-ARP USING A SYNTHETIC SCENE MADE OF TWO STANDARD RESOLUTION CARDS (UNIT MM)

Group	d	f	s	e	$d - d(\alpha^-)$	$d + d(\alpha^+)$
$\Omega_0 = 5.0m$	4250.00	99.65	34.88	2.01	3884.58	4691.31
	4750.00	99.89	34.96	2.25	4297.10	5309.62
	5000.00	100.00	35.00	2.37	4500.02	5624.68
	5300.00	100.11	35.04	2.51	4741.21	6008.11
	5800.00	100.28	35.10	2.76	5136.44	6660.45
$\Omega_0 = 6.0m$	5050.00	99.69	34.89	2.39	4540.54	5688.23
	5650.00	99.90	34.96	2.68	5018.67	6463.02
	6000.00	100.00	35.00	2.85	5292.28	6926.22
	6450.00	100.12	35.04	3.07	5638.49	7534.38
	7100.00	100.26	35.09	3.38	6127.80	8438.85
$\Omega_0 = 7.0m$	5500.00	99.61	34.86	2.61	4900.20	6267.11
	6500.00	99.89	34.96	3.09	5676.55	7602.89
	7000.00	100.00	35.00	3.34	6053.31	8297.70
	7600.00	100.11	35.04	3.63	6495.79	9156.51
	8800.00	100.29	35.10	4.21	7350.58	10961.41

[16] J. H. R. Chang, B. V. Vijaya Kumar, and A. C. Sankaranarayanan, "Towards multifocal displays with dense focal stacks," *ACM TOG*, vol. 37, no. 6, 2018.

[17] D. B. Murphy and M. W. Davidson, *Fundamentals of Light Microscopy and Electronic Imaging, Second Edition*. JOHN WILEY & SONS SINGAPORE PTE LTD., 2012.

[18] M. Martinez-Corral and B. Javidi, "Fundamentals of 3d imaging and displays: A tutorial on integral imaging, light-field, and plenoptic systems," *Advances in Optics and Photonics*, vol. 10, p. 512, 09 2018.

[19] M. Martínez-Corral and G. Saavedra, *Chapter 1 The Resolution Challenge in 3D Optical Microscopy*. Elsevier Science & Technology, 2009.

[20] Y. Chen, X. Jin, and Q. Dai, "Distance measurement based on light field geometry and ray tracing," *Optics Express*, vol. 25, no. 1, p. 59, 2017.

[21] C. Hahne, A. Aggoun, V. Velisavljevic, S. Fiebig, and M. Pesch, "Refocusing distance of a standard plenoptic camera," *Optics Express*, vol. 24, no. 19, p. 21521, 2016.

[22] M. Broxton, L. Grosenick, S. Yang, N. Cohen, A. Andalman, K. Deisseroth, and M. Levoy, "Wave optics theory and 3-D deconvolution for the light field microscope," *Optics Express*, vol. 21, no. 21, pp. 25418–

TABLE II
PARAMETERS OF THE IDEAL LIGHT FIELD CAMERA ARRAY CORRESPONDING TO THE EXPERIMENT OF FIGURE 8 (UNIT MM)

Array size	Refocus planes	d	f	s	e	$d - d(\alpha^-)$	$d + d(\alpha^+)$
9×9	$-d5$	1633.79	29.90	27.91	2.07	1579.19	1695.77
V	$-d4$	1695.81	29.92	27.92	2.15	1633.79	1762.72
30.46	$-d3$	1762.77	29.94	27.94	2.23	1695.81	1835.24
$r_x(\text{pix.})$	$-d2$	1835.29	29.96	27.96	2.33	1762.77	1914.02
512	$-d1$	1914.08	29.98	27.98	2.43	1835.29	1999.94
$r_y(\text{pix.})$	$-d0$	2000.00	30.00	28.00	2.54	1914.08	2094.00
512	$+d1$	2094.00	30.02	28.02	2.66	1999.94	2197.34
b	$+d2$	2197.34	30.04	28.04	2.79	2093.93	2311.49
113.14	$+d3$	2311.49	30.06	28.06	2.94	2197.27	2438.25
M'	$+d4$	2438.25	30.08	28.08	3.10	2311.41	2579.81
1.00	$+d5$	2579.81	30.10	28.09	3.29	2438.16	2738.94

25 439, 2013.

- [23] J. Jin, J. Hou, J. Chen, and S. Kwong, "Light field spatial super-resolution via deep combinatorial geometry embedding and structural consistency regularization," in *IEEE CVPR*, 2020, pp. 2257–2266.
- [24] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, "A dataset and evaluation methodology for depth estimation on 4d light fields," in *Asian Conference on Computer Vision*, 2017.
- [25] R. Hassen, Z. Wang, and M. M. A. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2798–2810, 2013.
- [26] B. Mildenhall, P. P. Srinivasan, R. Ortiz-Cayon, N. K. Kalantari, R. Ramamoorthi, R. Ng, and A. Kar, "Local light field fusion: Practical view synthesis with prescriptive sampling guidelines," *ACM SIGGRAPH*, vol. 38, no. 4, 2019.
- [27] J. Shi, X. Jiang, and C. Guillemot, "Learning fused pixel and feature-based view reconstructions for light fields," in *IEEE CVPR*, 2020, pp. 2552–2561.
- [28] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and N. Ren, "Nerf: Representing scenes as neural radiance fields for view synthesis," in *European Conference on Computer Vision*, 2020.

VI. APPENDIX

Let P_0 be the width of the conjugate virtual sensor in the object focal plane ω_0 . Figure 13 shows the proportional relation between P_0 and the sensor size s . Therefore, e can be derived under the pin-hole camera model assumption. We

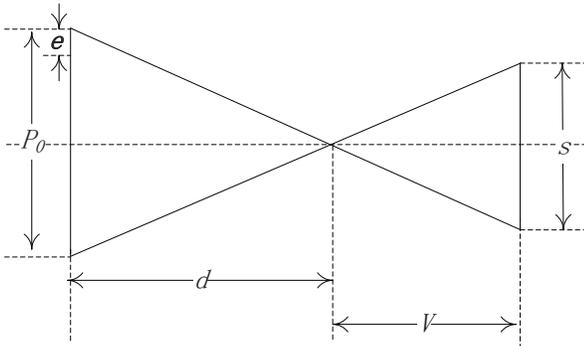


Fig. 13. The relation between P_0 , s and the theoretical shifting limit e of a pin-hole camera with a fixed image plane. Please note that the image plane distance V is a constant. Otherwise, the FoV will vary with different focal lengths f , so that P_0 can not be derived using Equation 12 below.

start the derivation from the simple geometry in Figure 13,

$$\frac{P_0}{d} = \frac{s}{V} \quad (10)$$

As mentioned in Section III C, we assume that the image plane distance V is a constant and that the FoV is fixed. Then V

can be simply inferred from the thin lens equation (Equation 5) as

$$V = \frac{(d - f)}{fd}. \quad (11)$$

By replacing V in Equation 10 by its expression in Equation 11, P_0 can be calculated as

$$P_0 = \frac{sd}{V} = \frac{s(d - f)}{f} \quad (12)$$

Then we can obtain e as

$$e = \frac{P_0}{\sqrt{r_x^2 + r_y^2}} = \frac{s(d - f)}{f\sqrt{r_x^2 + r_y^2}} \quad (13)$$

Further, the relation between $d(\alpha^+)$ and $d + d(\alpha^+)$ can be found by solving triangles in Figure 4, i.e.,

$$\frac{e}{(b/2)} = \frac{d(\alpha^+)}{d + d(\alpha^+)} \quad (14)$$

When using a magnifying lens, from image side to object side, the limit e will be multiplied by M' (the reciprocal of lateral magnification factor M), as

$$\frac{M'e}{(b/2)} = \frac{d(\alpha^+)}{d + d(\alpha^+)} \quad (15)$$

One can easily derive $d(\alpha^+)$ from Equation 14 as

$$d(\alpha^+) = \frac{dM'e}{b/2 - (M'e)} \quad (16)$$

In a similar way, $d(\alpha^-)$ can be derived as

$$d(\alpha^-) = \frac{dM'e}{b/2 + (M'e)} \quad (17)$$

Finally, the *Arp* model can be expressed as

$$\begin{aligned} \text{Arp}(LF_0, \Omega_0) &= \phi(d, f, b, s) \\ &= [d - d(\alpha^-), d + d(\alpha^+)] \\ &= \left[d \left(1 - \frac{M'e}{b/2 + M'e} \right), d \left(1 + \frac{M'e}{b/2 - M'e} \right) \right] \end{aligned} \quad (18)$$

which is the same as in Equation 6 in Section III C.