Continuous Refocusing for Integral Microscopy with Fourier Plane Recording

Sergio Moreschini[1], Gabriele Scrofani[2], Robert Bregovic[1], Genaro Saavedra[2] and Atanas Gotchev[1]

[1]*Laboratory of Signal Processing, Tampere University of Technology,* Tampere, Finland
Email: {sergio.moreschini, robert.bregovic, atanas.gotchev}@tut.fi

[2]*Department of Optics, University of Valencia,* Burjassot, Spain
Email: {gabriele.scrofani, genaro.saavedra}@uv.es

***Abstract*—Integral or light field imaging is an attractive approach in microscopy, as it allows to capture 3D samples in just one shot and explore them later through changing the focus on particular depth planes of interest. However, it requires a compromise between spatial and angular resolution on the 2D sensor recording the microscopic images. A particular setting called Fourier Integral Microscope (FIMic) allows maximizing the spatial resolution for the cost of reducing the angular one. In this work, we propose a technique, which aims at reconstructing the continuous light field from sparse FIMic measurements, thus providing the functionality of continuous refocus on any arbitrary depth plane. Our main tool is the densely-sampled light field reconstruction in shearlet domain specifically tailored for the case of FIMic. The experiments demonstrate that the implemented technique yields better results compared to refocusing sparsely-sampled data.**

***Index Terms*—refocusing, LF, microscopy, reconstruction, FIMic**

I. Introduction

In classical microscopy, a 3D sample is captured by taking numerous images, each at a different focal length (effectively slicing the sample at different depths), and recombining the images into a 3D structure. Since many images with narrow depth-of-field are required for achieving a good depth resolution, the process is slow. In practice, it is more desirable if the same information is captured at once. One way to accomplish this is by so-called integral, light field (LF) or plenoptic photography where in addition to the spatial information, the angular information (perspective) is captured, that in turn, can be used to make virtual refocused images at a desired depth. An example of such system is the integral camera introduced by Lippmann in 1908 [1]. Integral photography was brought to microscopes by Jang and Javidi [3]. The work was followed by Levoy et al. in [4] where microscopy was formally adapted to the concept of plenoptic function - a more formal description of the LF as introduced by Adelson and Bergen [2]. Following these initial steps, researchers focused on achieving the best possible result on an integral microscope by improving the resolution of the system. The main tools employed to achieve the goal were new deconvolution techniques, 4D interpolation and interpolation by time multiplexing [5]. With the aim of having a better control over the spatial and angular resolution, recently the so called Fourier Integral Microscope (FIMic) [8] has been proposed. It is used in this paper and will be described in more detail later on.

A key factor in LF microscopy is the available number of pixels which limits the overall amount of information that can be stored on the sensor thereby forcing a trade-off between the microscope's spatial and angular resolution. By refocusing the multiplexed LF data, one can reconstruct images (slices) corresponding to different focal planes [12]. For best results, one needs many high-resolution images captured from different perspectives (high angular resolution). However, an LF captured by FIMic can either have high spatial and low angular resolution or low spatial and high angular resolution.

In this paper, we demonstrate that samples with high spatial and low angular resolution are instrumental for computing the missing dense angular samples (views). This is done by performing a shearlet-based LF reconstruction [7]. Computationally, refocusing is done by shearing in LF domain, and we demonstrate the relation between the shearing factor and the re-focused depth for the microscope system under consideration. This relation can be used to map shearing factors into correct depths. Finally, the proposed refocusing approach is applied to a real object captured with the microscope as a demonstration that the proposed approach works well with real-world complex-depth scenes.

II. The Microscope

Figure 1 shows the optical arrangement of FIMic [8]. It is composed of five optical elements: a microscope objective (MO), two lenses L1 and L2 that act as a 4-f correlator, a field stop (FS) and a microlens array (MLA). The 4-f correlator conjugates the MLA with the aperture stop of the MO. Thus, the reference object plane (ROP) gets focused on the CCD behind each microlens. The FS limits the field of view (FOV) in such a way that the orthogonal elemental images (EIs) are tangent on the sensor plane. The resolution of the microscope is determined by lenses diffraction and by the sensor sampling rate. If the pixel size δ is chosen equal to the diffraction limit then the resolution power (in micrometers) and the depth of field (DOF) are given by [8]

$$\rho_{FIMic} = N\frac{\lambda}{2A_N} \tag{1}$$

ISBN 978-90-827970-1-5

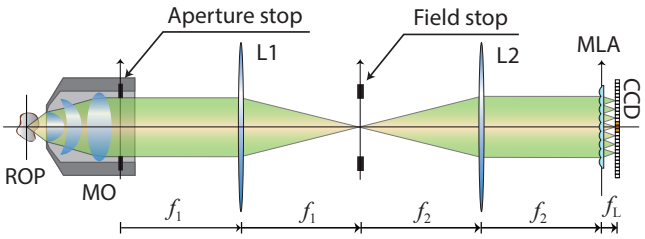


Fig. 1: Schematic of Far-field integral microscope (FIMic). The reference object plane (ROP) is directly focusing on each of the orthographic views or EIs.

and

$$DOF_{FIMic} = \frac{5}{4}\frac{\lambda}{2A_N^2}N^2, \tag{2}$$

where λ is the wavelength, A_N is the MO numerical aperture, and N is the number of microlenses that fit along the diameter of the MO aperture stop, which can be interpreted as the number of angular samples taken along one direction. As N increases, the resolution gets worse and the DOF gets better. The best spatial resolution is achieved for $N = 1$ for the complete lack of angular information and N, as low as $N = 3$, is needed in order to maintain full parallax.

The EIs captured by the microscope are indexed in rows and columns (i, j) as shown in Figure 2(a). Denote by $P(x_P, y_P, z_P)$ a point in space. Given t is the distance between the centers of two EIs, the disparity caused by the point P in mm on the sensor plane is

$$d_p = t\frac{f_L}{f_{MO}^2}\left(\frac{f_1}{f_2}\right)^2 z_P, \tag{3}$$

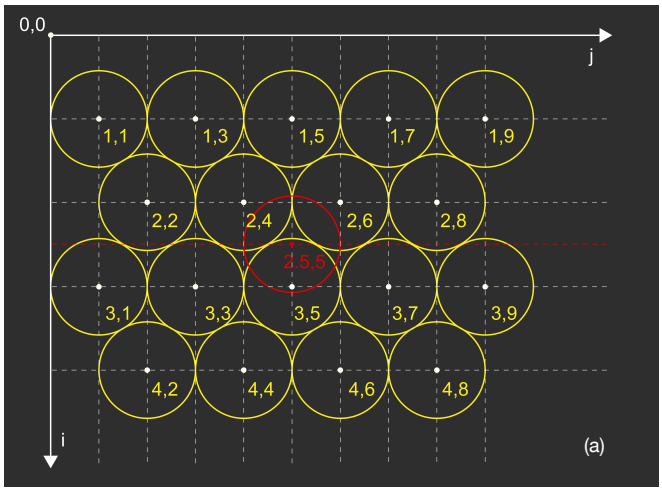
where f_L, f_{MO}, f_1 and f_2 are the focal lengths of the MLA, MO, and the first and second lenses, respectively. A central "virtual" elemental image to be used as a reference for disparity calculation, is denoted by $EI_{2.5,5}$, as depicted in Figure 2(a). For each $EI_{i,j}$, the coordinate system (u, v) is placed at the EI's centre, as shown in Figure 2(b). Eventually, the point $P(x_P, y_P, z_P)$ is mapped on $EI_{i,j}$ as $P_{i,j}(u, v)$

$$\begin{aligned} u_{p(i,j)} &= \frac{x_p M + d_p(2.5 - i)\cos(30^\circ)}{\delta}, \\ v_{p(i,j)} &= \frac{y_p M + d_p\frac{(5-j)}{2}}{\delta}, \end{aligned} \tag{4}$$

where $M = \frac{f_L}{f_{MO}}\frac{f_1}{f_2}$ is the magnification factor of the microscope. LF dataset $LF(u, v, i, j)$ is created by cropping and combing all EIs. When targeting high spatial resolution, the number of EIs is kept small with the aim to increase the angular resolution by means of view interpolation.

III. LF Reconstruction

The set of methods which aim at reconstructing the continuous plenoptic function is also known as Image-Based Rendering (IBR) [9], where each available pixel is considered as a multidimensional LF sample. In [10], a geometry-based



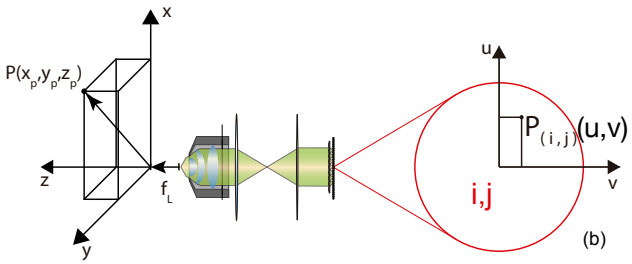


Fig. 2: (a) Indexing of the EIs. (b) Description of a point in space and central virtual image.

analysis for LF rendering has been conducted in order to define the minimum sampling rate, which is determined by the camera resolution and the scene disparity variation. In order to generate intermediate views without ghosting artefacts, one needs to reach a sampling rate where the disparity between neighbouring views is less than one pixel. This sampling condition is usually referred as dense sampling and the corresponding LF as the densely sampled LF (DSLF).

Consider a captured FIMic image, e.g. the one depicted in Figure 3(a), as sparsely sampled LF. By grouping rows for a fixed $i = i_0$, and for a specific axis value $u = u_0$, one gets a LF slice $E(v, j) = LF(u_0, v, i_o, j)$, which can be considered as Epi-polar plane image (EPI), as illustrated in Figure 3(b). The original EPI concept, as introduced by Bolles et al. in [11], considering camera motion, can also be applied to microscopy images where EI's different angular perspectives can be used for tracking points' variations through the EPI line slopes depending on their depth. In [7], EPI properties have been exploited in shearlet domain to reconstruct all views forming DSLF. Define d_H and d_V as the disparities over the horizontal and vertical axes respectively and define the horizontal disparity budget by the factor $\Delta d = \lceil d_{Hmax} - d_{Hmin} \rceil$. This factor determines the lines to be synthesized between each two given horizontal lines in EPI and consequently the number of scales and shearing factors in the shearlet transform [7]. The cone-adapted shearlet system which is the main tool used for EPI

ISBN 978-90-827970-1-5

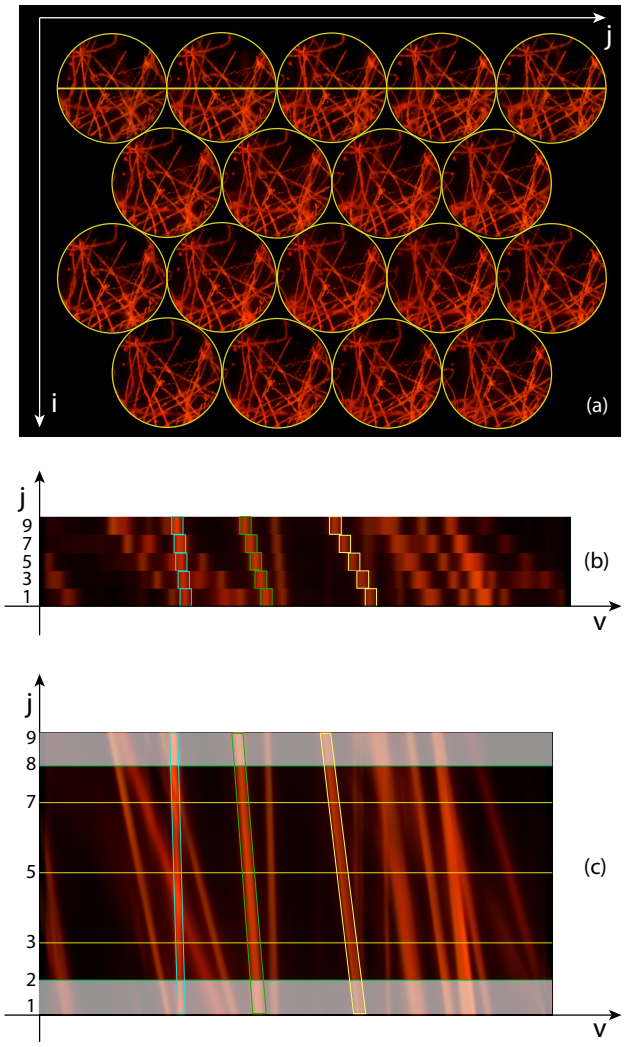


Fig. 3: Steps for Reconstruction. (a) Example of sparsely sampled LF captured with FIMic; central line marked in yellow. (b) EPI for central line of sparsely sampled LF. (c) EPI for central line of reconstructed DSLF.

sparsification is composed of two cones regions and a low-pass region. The elements of the shearlet system, defined as atoms, are generated from a translation of a scaling function and translation, shearing and scaling of two shearlets. The atoms at different scales are constructed dyadically and the optimal number of scales generated is equal to

$$J = \lceil \log_2 \Delta d \rceil. \tag{5}$$

A higher value would guarantee a better result but would also increase the computational load. A lower value would produce a substantial reduction of the reconstruction quality as the number of shearing atoms would not be enough to have support for the whole region where the EPI is defined in the frequency domain. A DSLF EPI with all missing rows reconstructed is shown in Figure 3(c).

The process of reconstructing DSLF from EIs is performed in two steps, where horizontal images are reconstructed first, followed by reconstruction in vertical direction (see Figure 4 for illustration). In both steps, the reconstruction algorithm described in [7] has been used.

The criterion for dense sampling requires the minimum amount of novel views to be generated between two adjacent views to be equal to $\lceil \Delta d_H \rceil - 1$, denoted as R_H. However, to be able to generate views matching the samples vertically, the number of images created between two existing views needs to be an odd number. This is illustrated in Figure 4(a) - e.g. after reconstructing the required horizontal views in the first row (images $1,1$ thorough $1,9$), one needs to reconstruct images that match vertically the images from the second row, e.g. $1,2$. This can only be achieved if R_H is odd. Furthermore, from refocusing viewpoint, it is better to have a large number of images. Therefore we opt to use a higher upsampling factor than required by the DSLF factor, which leads to $R_H = (2 \cdot \lceil \Delta d_H \rceil) - 1$ images. In the FIMic arrangement, different rows are composed of different number of samples. Therefore, after performing all horizontal reconstructions, for the odd rows (i.e. first and third) some readjustment is required. As depicted in Figure 4(a), some images are dropped. Their number is equal to $2l_H$, where $l_H = \frac{(R_H+1)}{2}$. The new dataset will then comprehend in the horizontal direction a number of $n = 4 + 3 \cdot R_H$ images starting from $j = 2$ to $j = 8$.

The geometry of the microlens array determines the maximum disparity between adjacent sparse samples on the vertical axis: $\Delta d_V = \lceil \Delta d_H \cdot \cos(30°) \rceil$, which requires generating $R_V = (2 \cdot \Delta d_V) - 1$ novel views between two adjacent images in vertical direction for any available column as depicted in Figure 4(b).

IV. Refocusing

In [12], it has been discussed that the dynamic representation of LFs can be achieved by performing a shear transformation in the ray space. In particular, any point feature in the scene corresponds to a particular line in the EPI. Therefore, each slope in the EPI corresponds to the depth of each single feature. The elements which are characterized by a vertical slope in the ray space are those which are in focus, the slopes with positive values describe objects which are far away from the focus plane, while if the slope is negative the represented element is closer than the focal plane. Refocus at a specific depth (Z_p) can be done by performing a specific shear in the ray space. The amount of shearing can be calculated as:

$$S = \frac{Z_p}{R+1} \frac{f_L}{f_{MO}^2} \left(\frac{f_1}{f_2} \right)^2 \frac{p}{\delta}, \tag{6}$$

where R is the number of images betweent EIs, p is the pitch of the MLA and δ is the pixel size of the sensor. The relation between the shearing and disparity is driven by Eqs. (6) and (3). DSLF is a particularly suitable representation for performing refocusing through shearing as it provides enough samples for ray interpolation.

ISBN 978-90-827970-1-5

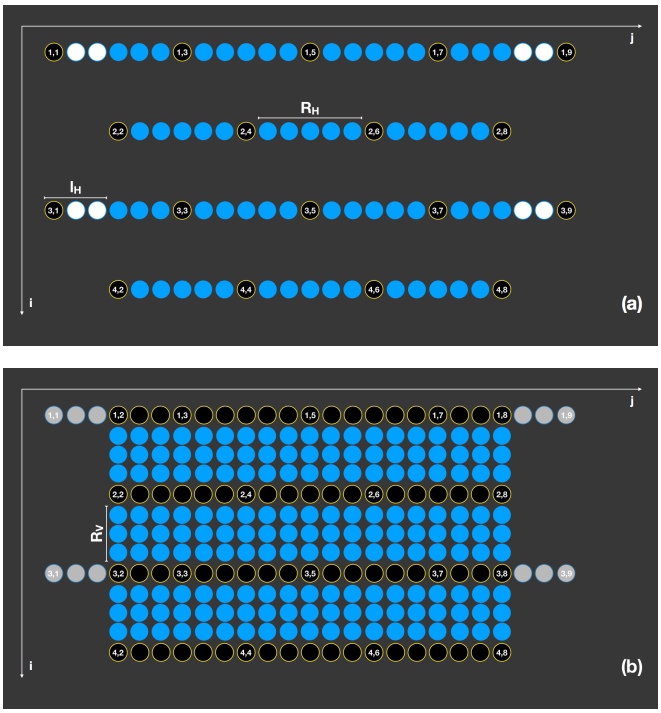


Fig. 4: From sparsely sampled LF to densely sampled LF. (a) Horizontal parallax reconstruction. (b) Vertical parallax reconstruction.

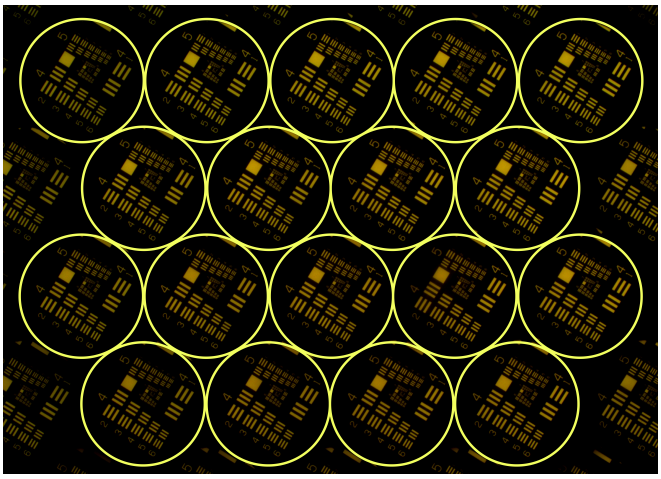


Fig. 5: Sparsely-sampled data for USAF test chart captured at 100 μm from the ROP.

V. Experiments

We present two sets of experiments using data captured by FIMic. In the first set, an USAF test chart was captured at different depths. This provides several sparsely sampled LFs with different disparities. When capturing a sparse sample, the test chart is placed at the focus point then, once captured the first image ($z_{min} = 0\ \mu m$), with the help of micrometer axial displacer, the successive images are captured at difference steps of 100 micrometers until reaching the limit of the DOF ($z_{max} = 400\ \mu m$). An example of the captured image

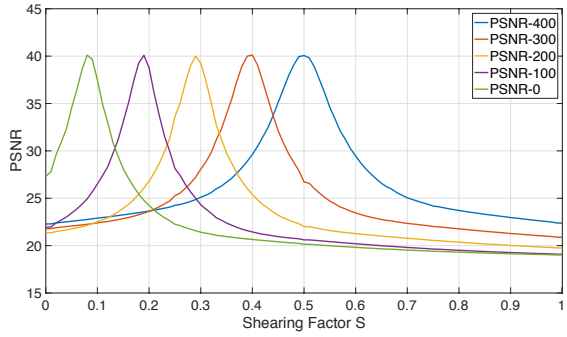


Fig. 6: PSNR changes over different frames for the different densely sample LFs.

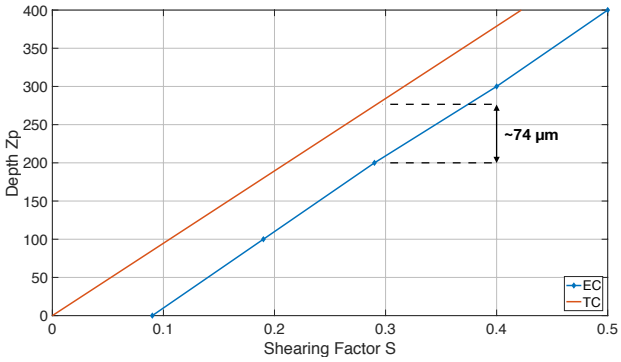


Fig. 7: The experimental curve (EC in blue) increase linearly with the amount of pixels shifted with the shearing according with the theoretical curve (TC in red), from which it differs by an offset of 74 μm.

from which we extrapolate the different views of dimension 455x455 pixels is depicted in Figure 5.

Figure 6 shows the Peak Signal to Noise Ratio (PSNR) between refocused images and one of the existing images, namely $EI_{3,5}$ for different depths of the USAF test chart. The peak for each dataset is reached when the shearing factor matches the captured depth. One is to note that the position of the test chart has an offset; this is a systematic error caused by misalignments in the FiMic building process. This offset spreads among all samples shifting the peaks of roughly 74 μm. To take this offset into account, in the formulas one has to substitute $Z_p = Z_{Exp} - Z_{OS}$, where Z_p is the real depth, Z_{Exp} is the experimental position in depth and Z_{OS} is the offset in depth of the system. The experimental results show a linear behaviour as predicted in Eq. (3) and also illustrated in Figure 7. The result shown before have been compared with the standard "shift and sum" ($S\&S$) refocusing technique usually employed for this specific microscope data [13]. The comparisons between the DSLF Refocusing (DSLFR) technique and the $S\&S$ at depth $Z_R = 200$ micrometers is depicted in Figure 8, where an improvement of $\approx 2.5\ dB$ is seen.

The second set of experiments has been performed on volumetric sample of cotton fibre captured by FIMic. A visual

ISBN 978-90-827970-1-5 © EURASIP 2018

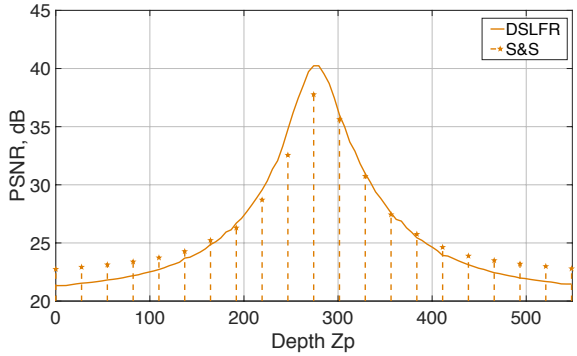


Fig. 8: Comparison of the "shift and sum" (S&S) and the DSLF Refocusing (DSLFR) technique.

comparison of the two approaches for two different S values is given in Figure 9. One can notice that DSLF refocusing technique, shown in Figures 9 (a) and (c), resolves aliasing problems otherwise present in Figures 9 (b) and (d). In the supplement video (https://youtu.be/PGupKTwJ6dE), one can explore the "continuity" of the DSLFR approach compared against the "shift and sum" which refocuses only on a limited number of planes.

VI. Conclusions

3D light field microscopy allows to capture the scene of interest at once and to explore it later by refocusing at different depths. In this process, maintaining high spatial and angular resolution is of prime importance. However, these two work in opposite directions. Furthermore, technological limitations of sensors and optics impose constraints, which have to be overcome by computational means. We have demonstrated that the FIMic optical setting, which offers high spatial resolution for the price of reduced angular resolution, can be successfully equipped with a computational imaging tool, which effectively reconstructs the desired number of perspective directions thus allowing to maintain continuous refocus at any arbitrary depth plane of interest. The tool utilizes the shearlet-domain light field reconstruction technique, which was modified for the case of FIMic. It is flexible in determining the number of scales and directions, which can be manipulated to get an over-sampled light field. This can be regarded as an effective directional interpolation, which is instrumental for the continuous refocusing operation. Our experiments with a resolution chart and a real-world object demonstrated the high quality of the refocused images compared against the standard technique used in this type of microscopy.

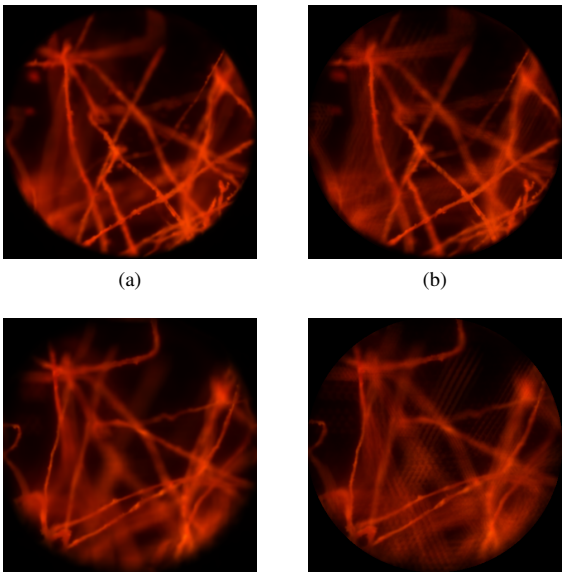


Fig. 9: Visual comparison of different refocusing techniques. (a) DSLFR at $Z_p = 98.67\ \mu m$. (b) "$S\&S$" at $Z_p = 98.67\ \mu m$. (c) DSLFR at $Z_p = 345.37\ \mu m$. (d) "$S\&S$" at $Z_p = 345.37\ \mu m$.

References

[1] M. G. Lippmann, "Epreuves reversibles donnant la sensation du relief," *J. de Phys. Theor. Appl.,* vol. 7, no. 7, pp. 821-825, Nov. 1908.

[2] E. H. Adelson and J. Y. Wang. "Single lens stereo with a plenoptic camera," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 14 no. 2, pp. 99-106, Feb. 1992.

[3] J. S. Jang and B. Javidi, "Three-dimensional integral imaging of micro-objects," *Opt. Lett.,* vol. 29, no. 11, pp. 1230-1232, Jun. 2004.

[4] M. Levoy, R. Ng, A. Adams, M. Footer and M. Horowitz, "Light Field microscopy," *ACM Trans. Graph,* vol. 25, no.3, pp. 924-934, Jul. 2006.

[5] A. Llavador, E. Sánchez-Ortiga, J. C. Barreiro, G. Saavedra and M. Martínez-Corral. "Resolution enhancement in integral microscopy by physical interpolation," *Biomed. Opt. Express,* vol. 6, no 8, pp. 2854-2863, Jul. 2015

[6] A. Llavador, J. Sola-Pikabea, G. Saavedra, B. Javidi and M. Martínez-Corral, "Resolution improvements in integral microscopy with Fourier plane recording," *Opt. Express,* vol. 24, no. 18, pp. 20792-20798, Aug. 2016.

[7] S. Vagharshakyan, R. Bregovic and A. Gotchev. "Light Field reconstruction using Shearlet transform," *IEEE Pattern Anal. Mach. Intell.,* vol. 40, no. 1, pp.133-147, Jan. 2018.

[8] G. Scrofani, J. Sola-Pikabea, A. Llavador, E. Sánchez-Ortiga, J. C. Barreiro, G. Saavedra and M. Martínez-Corral, "FIMic: design for ultimate 3D-integral microscopy of in-vivo biological samples," *Biomed. Opt. Express,* vol. 9, no. 1, pp. 335-346, Jan. 2018.

[9] H. Shum, S. Chan and S. Kang, *Image-Based Rendering*. New York: Springer, 2007.

[10] Z. Lin and H.-Y. Shum. "A geometric analysis of light field rendering," *Int. J. Comput. Vis.,* vol. 58, no. 2, pp. 121-138, Jul. 2004.

[11] R. Bolles, H. Baker and D. Marimont. "Epipolar-plane image analysis: An approach to determining structure from motion," *Int. J. Comput. Vis.,* vol. 1, no. 1, pp. 7-55, Mar. 1987.

[12] A. Isaksen, L. McMillan and S. J. Gortler. "Dynamically reparameterized light fields," *Proc. 27th Annu. Conf. Compu. Graphics Interactive Techn.,* 2000, pp. 297-306.

[13] R Ng, M Levoy, M Brédif, G Duval, M Horowitz and P Hanrahan. "Light field photography with a hand-held plenoptic camera." Computer Science Technical Report CSTR, Stanford University, Stanford, CA, USA, 2005.

ISBN 978-90-827970-1-5