A Depth Sensor Based on Transient Property of Liquid Crystal Lens

Haifeng XIAO, Zhiqiang LIU, Baolin TAN, and Mao YE*

School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

*Corresponding author: Mao YE E-mail: mao_ye@uestc.edu.cn

Abstract: The property of maintaining the lens state of the liquid crystal (LC) lens during the switching between positive and negative lens states is made use of in the fast acquirement of multi-focus images without magnification change. A depth from focus (DFF) pipeline that can generate a low-error depth map and an all-in-focus image is proposed. The depth of the scene is then obtained via DFF pipeline from the captured images. The depth sensor proposed in this paper has the advantages of simple structure, low cost, and long service life.

Keywords: Liquid crystal lens; transient property; depth from focus; depth sensor

Citation: Haifeng XIAO, Zhiqiang LIU, Baolin TAN, and Mao YE, "A Depth Sensor Based on Transient Property of Liquid Crystal Lens," *Photonic Sensors*, 2023, 13(2): 230230.

1. Introduction

With the continuous development of computer vision and sensor technology, more and more depth measurement methods have been proposed. Mainstream depth measurement methods include active depth measurement methods such as time of flight (TOF) [1] and structured light [2] and passive depth measurement methods such as binocular stereo vision [3], depth from defocus (DFD) [4], and depth from focus (DFF) [5]. The active depth measurement method requires projecting laser or encoded light into the scene and obtaining the depth by collecting and analyzing the reflected optical signal, and the accuracy of such methods is relatively high, but it is sensitive to the ambient light. The passive depth measurement method needs to acquire the scene images and calculate the depth by analyzing the parallax, sharpness, or other information contained in the images. It relies on the texture information of the scene. In the passive method, binocular stereo vision obtains depth by calculating image parallax and is now widely used in smartphone cameras [6], intelligent robots [7], truth detection [8], and other fields. However, camera calibration and image matching algorithms are time-consuming. DFD and DFF both calculate the depth by the sharpness of the pixels. DFD obtains the depth from two out-of-focus images by establishing the relationship between the size of the circle of blur and the depth. It has been applied in the automatic focusing of cameras [9]. DFF needs multi-focus images to estimate the depth, and the depth resolution is directly related to the number of images. Tens even hundreds of images are often used for depth estimation. Lenses are moved to capture a multi-focus image stack. However, mechanical movements of lenses will reduce the

Received: 18 May 2022 / Revised: 21 July 2022

DOI: 10.1007/s13320-022-0669-2
Article type: Regular

accuracy and the life of the system. Kawamura *et al.* [10] have proposed a DFF method using a liquid crystal (LC) lens that achieved focusing by varying the driving voltages. The voltages were adjusted sequentially to obtain a multi-focus image stack. The method was later followed by Li *et al*. [11] and Emberger *et al*. [12]. The LC lens is slow to respond to voltages. It is time-consuming to collect the image stack necessary for DFF. The method therefore cannot be considered effective. We have found that during the switching between the positive and negative lens states of the LC lens [13], the root mean square (RMS) aberration was below 0.1 wavelength [14]; lens property remains, while the optical power changes with time. Multi-focus image stack can then be captured during the switching.

In this paper, we report a depth acquisition method via DFF using an LC lens. The depth is extracted from the multi-focus image stack captured during the switching, which usually takes only several seconds of the LC lens from the positive state to the negative state. We propose a depth acquisition pipeline to improve the performance of Tseng's model [15] on DFF, and the low-error depth map is obtained. In the traditional DFF method [4, 5, 16], the CMOS (complementary metal oxide semiconductor) sensor or lens position is continuously changed by a stepper motor to obtain a multi-focus image sequence, and the magnification between the acquired images is inconsistent. The LC lens as a system diaphragm can ensure the consistency of magnification between images. In the DFF method using an LC lens proposed in [10–12], the sequential voltage regulation method is adopted, and the image acquisition time is positively correlated with the number of images. Using the transient property of the LC lens, the DFF method reported in this paper makes full use of the time in the process of state switching to collect dense multi-focus scene images and obtain more accurate depth data than the DFD method reported in [17], which takes almost the same time to collect the two images for the calculation. Therefore, introducing the transient property of the LC lens into the DFF method can effectively balance the cost, speed, and accuracy of depth extraction.

2. Liquid crystal lens

The LC lens structure [13] is shown in Fig. 1. The main body is three glass substrates plated with electrodes, and the thickness of the glass substrates is 0.7 mm. There are two ITO (indium tin oxide) electrodes and one Cr electrode with a round hole of 2.0 mm diameter in the center. The bottom surface of the glass substrate 2 and the surface of the ITO electrode of the glass substrate 3 have a layer of polyimide film, and their rubbing directions are anti-parallel. The LC (HTW109100-100 from JIANGSU HECHENG DISPLAY TECHNOLOGY CO. LTD, Δn = 0.259, n_e = 1.764, and n_o = 1.505) of 30 μm thickness is sandwiched between two glass substrates.

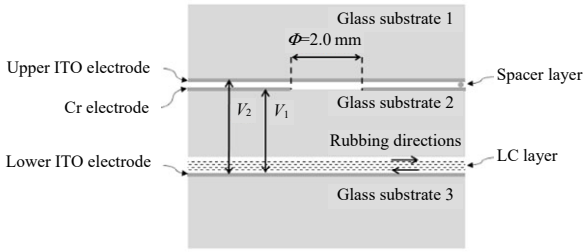


Fig. 1 Structure diagram of the LC lens.

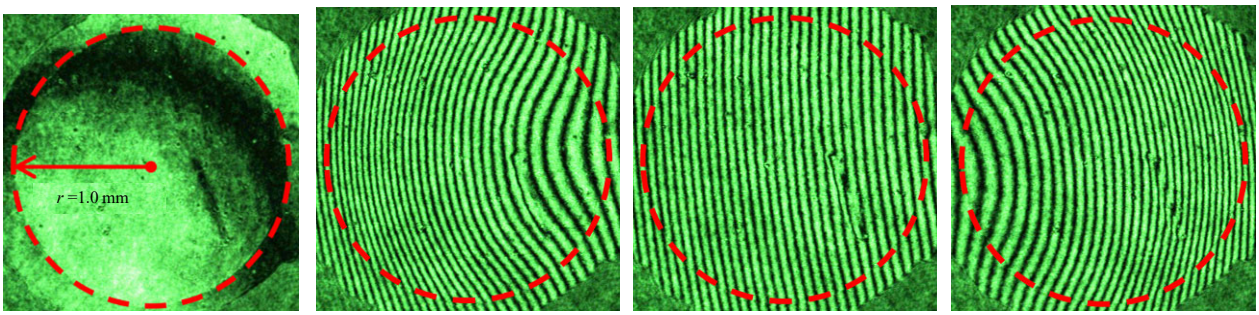
The LC cell is driven by two voltages V_1 and V_2. The lens properties of maximum positive and maximum negative states are shown in Table 1.

Table 1 Maximum optical power and RMS aberration.

LC lens state	V_1 (V_{RMS})	V_2 (V_{RMS})	P (m^{-1})	RMS aberr. (λ)
Positive	61	10	5.6	0.087
Negative	10	100	–4.3	0.088

A Mach-Zehnder interferometer with a laser beam of 532 nm wavelength is used to measure the lens properties. Figure 2 shows the interference

fringe pattern acquired during the state switching. The wavefront of the light beam emitted from the LC lens can be reconstructed by interference fringes. Then, the wavefront is fitted into Zernike polynomial to obtain the optical power and RMS aberration of the LC lens. Figure 3 shows the optical power and RMS aberration changing with time during the state switching. It can be seen that the aberration is low and the LC cell is considered to be a lens in the state change.



(a) (b) (c) (d)

Fig. 2 Interference fringe patterns collected during the state switching of the LC lens. The red annular dashed line is the boundary of the effective aperture of the LC lens: (a) a reference image used to determine the effective aperture of the LC lens, (b) the LC lens is in positive state, (c) the power of the LC lens is approximately equal to 0, and (d) the LC lens is in the negative state.

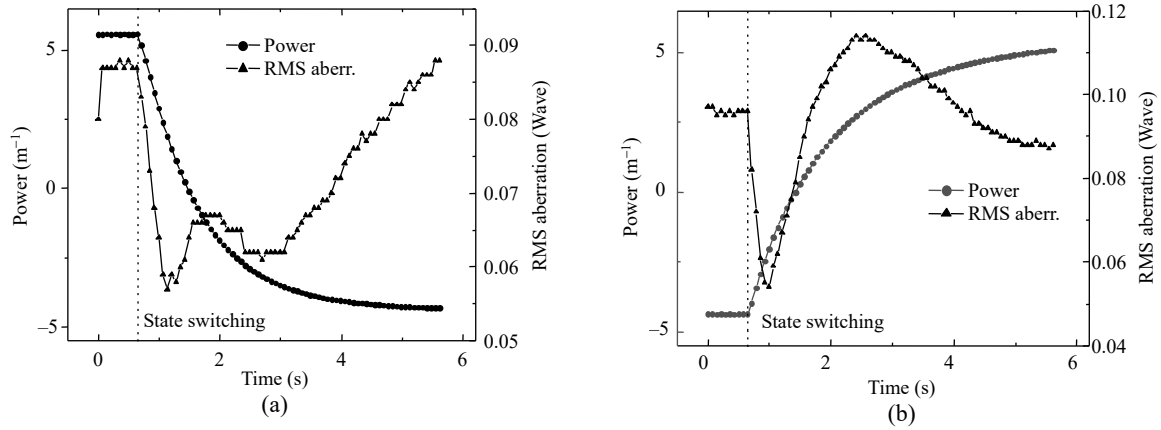


Fig. 3 LC lens property corresponds to two switching modes respectively: (a) the positive to negative lens states and (b) the negative to positive lens states. The LC lens starts state switching when the 10th picture is acquired.

3. LC lens imaging system

As shown in Fig. 4(a), the imaging system is composed of a CMOS (complementary metal oxide semiconductor) sensor, a main lens of 25 mm focal length, and an LC lens. The LC lens is placed in front of the main lens with a distance of 5.5 mm. The aperture of the LC lens is the diaphragm of the system. The CMOS sensor model is MD-50T (the resolution is 640×480), and its pixel size is 2.2 μm. The image distance v does not change. When the focal length of the LC lens changes, objects of different distances u are focused, and the magnification of the system remains constant [14].

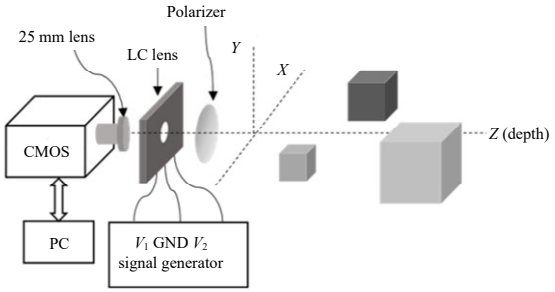
The object distance u, the image distance v, and the optical power P satisfy the Gaussian formula

$$\frac{1}{u}+\frac{1}{v}=P \tag{1}$$

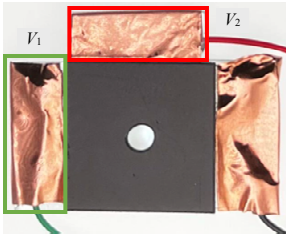
$$P=P_{\mathrm{LC}}+P_{\mathrm{main}}-dP_{\mathrm{LC}}P_{\mathrm{main}} \tag{2}$$

where P_{main} and P_{LC} are the powers of the main lens and the LC lens, respectively, and d is the distance between the main lens and the LC lens.

In DFF, the depth resolution is determined by the depth of field and the object distance interval of the



(a)



(b)

Fig. 4 LC lens imaging system: (a) configuration of the LC lens imaging system and (b) LC lens with a mask.

neighboring frames. The depth of field is defined by

$$\mathrm{DOF} = \frac{2F\delta u^2 P^2}{1 - F^2\delta^2 u^2 P^4} \tag{3}$$

where F is the aperture value and δ is the diameter of the blur circle; the pixel size of 2.2 μm is taken as the diameter. In this work, we would like to limit the error of the depth measurement within 1.0 cm. The system initially focuses on the object at 16.4 cm distance when the LC lens does not work. When the optical power of the LC lens changes from 5.6 m^{-1} to –4.3 m^{-1}, objects of the distance from 9.6 cm to 36.6 cm are focused sequentially, and the depth of field changes from 0.1 cm to 1.0 cm, as shown in Fig. 5.

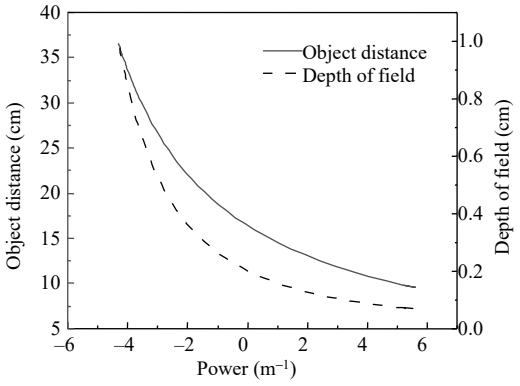


Fig. 5 Change of object distance and depth of field with power.

During state switching, the changes of the distance of focused objects with time obtained from Figs. 3 and 5 are shown in Fig. 6. It can be seen that it takes 3.3 s and 1.7 s approximately for the LC lens to switch from the positive state to the negative state and vice versa, respectively. The time needed to collect the images for the depth calculation is only several seconds and is far less than that for the method reported in [10–12]. More uniform distance change is observed during switching from positive to negative lens state and frames are captured in the process for depth calculation. Figure 7 shows images captured during state switching.

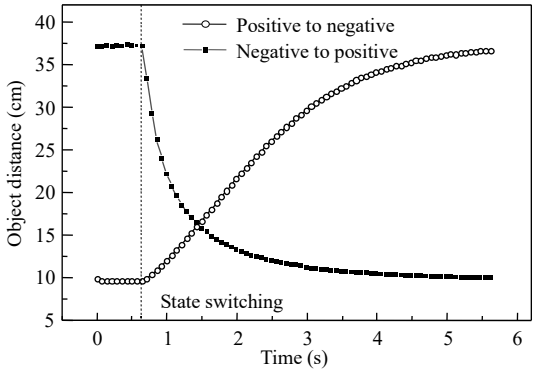


Fig. 6 Change of object distance with time.

(a)

(b)

Fig. 7 Images captured during positive to negative states switching: (a) Scene 1: card, from left to right are the images with focus at 10 cm, 20 cm, and 30 cm and (b) Scene 2: wallpaper, from left to right are the images focus at 10 cm, 12 cm, and 14 cm.

4. Depth from focus

The idea of DFF is to scan the whole scene with the focusing plane along the optical axis and estimate the depth using the imaging system parameters when the object is in focus. The DFF

includes focus measure and depth acquisition, which are described in turn below.

The focus measure is a measure of how much a pixel is in focus. In this work, the gradient operator is used to evaluate the degree of pixel focus. When the influence of noise is not considered, the gradient value of the pixel in the focus state reaches the maximum. Pertuz *et al.* [18] evaluated the performance of most common gradient operators, and the modified Laplacian (MLAP) operator proposed by Nayar *et al.* [5] performed well.

In the ith image I^i ($1 \leqslant i \leqslant K$), the degree of focus $FM^i(x, y)$ of the pixel located at (x, y) can be calculated as follows:

$$FM^i(x, y) = \frac{1}{(2r+1)^2} \sum_{x=x-r}^{x+r} \sum_{y=y-r}^{y+r} ML^i(x, y),$$
$$T_1 \leqslant ML^i(x, y) \leqslant T_2 \tag{4}$$

where $2r+1$ is the window size, and the threshold T is to reduce the impact of noise and exposure abnormalities. When the value of $ML^i(x, y)$ exceeds this range, it is set to 0. The values of r, T_1, and T_2 are 1, 20, and 230, respectively. $ML^i(x, y)$ is the gradient of the pixel (x, y) in the image I^i, which is defined as follows:

$$ML^i(x, y) = \left|2I^i(x, y) - I^i(x-1, y) - I^i(x+1, y)\right| + \left|2I^i(x, y) - I^i(x, y-1) - I^i(x, y+1)\right|. \tag{5}$$

Obtain the gradient maps FM^i ($1 \leqslant i \leqslant K$), and then calculate the initial depth map D according to the winner-take-all rule (WTA) [19]. WTA rule is expressed as

$$j = \arg\max_i \left[FM^i(x, y)\right] \tag{6}$$

$$D(x, y) = u(j) \tag{7}$$

where $u(j)$ is the object distance when the image I^j is acquired, which is obtained from the acquisition time of the image I^j and Fig. 6.

However, the focus measure cannot correctly reflect the focusing degrees of the weak texture and the abnormal exposure regions, and there is a lot of noise in the initial depth map D generated by the traditional method [5, 20]. The proposed depth acquisition pipeline is shown in Fig. 8.

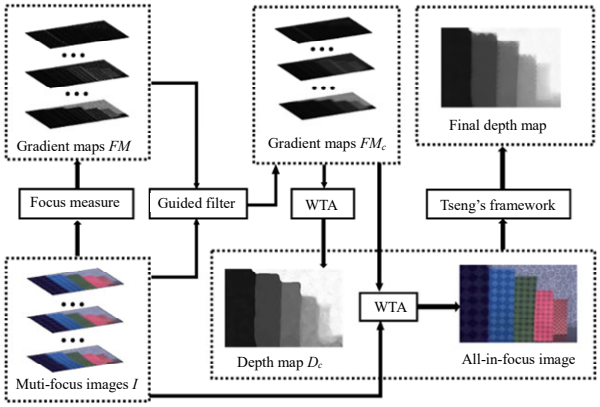


Fig. 8 Depth acquisition pipeline proposed in this work, with the arrow as input or output.

In the initial depth map D, the depth in the richly textured areas of the scene is more plausible, while the depth in the untextured areas and abnormal exposure areas tend to have errors. In order to propagate depth from regions with high reliability to regions with low reliability, a cost aggregation algorithm based on guided filter is used in this work [21, 22]. The image I^i is used as the guide map, and the gradient map FM^i is used as the processed map input into the guided filter. Using the edge-preserving property and color linear model of guided filter, the gradient of the object edge can be propagated to the non-textured area inside the object. The output is the gradient map FM_c^i, and then the depth map D_c can be obtained according to the WTA rule. We use Tseng's optimization framework [15] for post-processing. In addition to the depth map D_c, the framework needs an all-in-focus (AIF) image of the scene. The AIF image is obtained from the gradient maps FM_c^i ($1 \leqslant i \leqslant K$) and the multi-focus images I^i ($1 \leqslant i \leqslant K$), namely:

$$j = \arg\max_i \left[FM_c^i(x, y)\right] \tag{8}$$

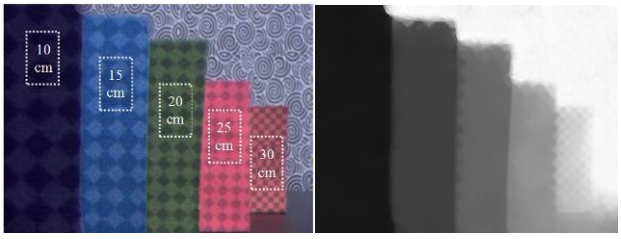
$$\mathrm{AIF}(x, y) = I^j(x, y). \tag{9}$$

After cost aggregation, the noise in depth map D_c is much less than the initial depth map D, and more correct depth value can be used by Tseng's framework. In other words, our pipeline is less reliant on the information in the all-in-focus image, which inhibits the replication of textures from the all-in-focus image into the final depth map.

Using a PC with Intel i5-6300HQ CPU and an 8G operating memory, it takes 150 seconds approximately for the depth extraction from 80 multi-focus images. The guided filter takes more than 80% of the time. We are considering using a more efficient cost aggregation algorithm to reduce time consumption.

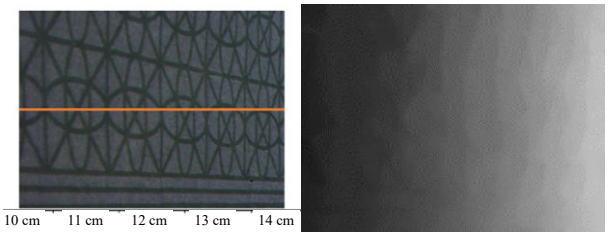
5. Experimental results

Figures 9(b) and 10(b) show depth maps, in which the color from black to white represents the distance from the nearest to the farthest. For Scene 1, the average depth within the region of interest (ROI) inside the cards is selected as the calculated value, and the ROI of each card is located inside the white box in Fig. 9(a). The depth inside the card maintains the consistency, and the depth error is within 4.3%. Table 2 shows the error of the depth map of Scene 1. For Scene 2, the depth value on a profile with a width of 10 pixels and length of 640 pixels is used as the calculated value, as shown in the orange line in Fig. 10(a). Scene 2 is a scene with smooth depth change. The depth value on the profile hardly jumps and reverses. Table 3 shows the errors of depth map. The average depth error of Scene 2 is 2.1%, and the maximum depth error is controlled at 4.1%.



(a) (b)

Fig. 9 Calculation result of scene 1: (a) all-in-focus image and (b) depth map.



(a) (b)

Fig. 10 Calculation result of scene 2: (a) all-in-focus image and (b) depth map.

Table 2 Depth error of Scene 1.

Object distance (cm)	Calculated value (cm)	Relative error (%)
10	10.43	4.3
15	14.56	2.9
20	19.14	4.3
25	24.25	3.0
30	30.56	1.2

Table 3 Depth error of Scene 2.

Relative error	Maximum	Mean
Value	4.1%	2.1%

6. Conclusions

We find that the LC lens maintains the lens property during state switching, and this paper proposes to apply the finding to the DFF method. By switching the state of the LC lens, the image stack required for depth calculation is obtained within seconds and the error of depth map calculated from it is less than 4.3%. The proposed DFF method is a suitable choice for depth measurement systems that require low cost and miniaturization. We are considering using a large-aperture LC lens with a new electrode structure to measure the scene depth in a wider range.

Acknowledgment

This study was supported by Sichuan Science and Technology Programs (Grant No. 2021YJ0102).

References

[1] M. Hansard, S. Lee, O. Choi, and R. Horaud, "Time-of-flight cameras: principles, methods and applications," *Springer Science & Business Media*, 2012.

[2] P. J. Besl, "Active optical range imaging sensors," *Advances in Machine Vision*, New York: Springer, 1989, pp. 1–63.

[3] W. E. L. Grimson and S. Brenner, "A computer implementation of a theory of human stereo vision," *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 1981, 292(1058): 217–253.

[4] Y. Xiong and S. A. Shafer, "Depth from focusing and defocusing," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, USA, 1993, pp. 68–73.

[5] S. K. Nayar and Y. Nakagawa, "Shape from focus: An effective approach for rough surface," in *Proceedings of IEEE International Conference on Robotics and Automation*, USA, 1990, pp. 218–225.

[6] F. Oniga, A. Trif, and S. Nedevschi, "Stereovision for obstacle detection on smart mobile devices: First results," in 1*6th International IEEE Conference on Intelligent Transportation Systems*, Netherlands, 2013, pp. 342–347.

[7] K. Okada, M. Inaba, and H. Inoue, "Integration of real-time binocular stereo vision and whole body information for dynamic walking navigation of humanoid robot," in *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Japan, 2003, pp. 131–136.

[8] F. Oniga and S. Nedevschi, "Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection," *IEEE Transactions on Vehicular Technology*, 2009, 59(3): 1172–1182.

[9] X. Zhang, Z. Liu, M. Jiang, and M. Chang, "Fast and accurate auto-focusing algorithm based on the combination of depth from focus and improved depth from defocus," *Optics Express*, 2014, 22(25): 31237–31247.

[10] M. Kawamura, S. Ishikuro, and S. Sato, "Imaging system for determining depth mapping properties by using a liquid crystal lens," in *IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society*, Japan, 2015, pp. 001966–001969.

[11] L. Hui, P. Fan, W. Yuntao, Z. Yanduo, and X. Xiaolin, "Depth map sensor based on optical doped lens with multi-walled carbon nanotubes of liquid crystal," *Applied Optics*, 2016, 55(1): 140–147.

[12] S. Emberger, L. Alacoque, A. Dupret, N. Fraval, and J. L. de Bougrenet de la Tocnaye, "Evaluation of the key design parameters of liquid crystal tunable lenses for depth-from-focus algorithm," *Applied Optics*, 2018, 57(1): 85–91.

[13] B. Wang, M. Ye, and S. Sato, "Liquid crystal lens with focal length variable from negative to positive values," *IEEE Photonics Technology Letters*, 2005, 18(1): 79–81.

[14] Y. Bai, X. Chen, J. Ma, J. Zeng, and M. Ye, "Transient property of liquid crystal lens and its application in extended depth of field imaging," *Optics Communications*, 2020, 473: 125974.

[15] C. Y. Tseng and S. J. Wang, "Shape-from-focus depth reconstruction with a spatial consistency model," *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, 12(24): 2063–2076.

[16] M. Moeller, M. Benning, C. Schonlieb, and D. Cremers, "Variational depth from focus reconstruction," *IEEE Transactions on Image Processing*, 2015, 24(12): 5369–5378.

[17] M. Ye, X. Chen, Q. Li, J. Zeng, and S. Yu, "Depth from defocus measurement method based on liquid crystal lens," *Optics Express*, 2018, 26(22): 28413–28420.

[18] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, 2013, 46(5): 1415–1432.

[19] A. R. Mansouri and J. Konrad, "Bayesian winner-take-all reconstruction of intermediate views from stereoscopic images," *IEEE Transactions on Image Processing*, 2000, 9(10): 1710–1722.

[20] M. Subbarao and T. Choi, "Accurate recovery of three-dimensional shape from image focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, 17(3): 266–274.

[21] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2012, 35(2): 504–511.

[22] H. G. Jeon, J. Surh, S. Im, and I. S. Kweon, "Ring difference filter for fast and noise robust depth from focus," *IEEE Transactions on Image Processing*, 2019, 29: 1045–1060.