IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 16, 2020, accepted March 31, 2020, date of publication April 6, 2020, date of current version April 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2986057

Real Scene Pickup Method of Elemental Image Array Based on Convergent Camera Array

MIN GUO[1], YUANZHI LYU[2], AND SHIGANG WANG[3]

[1]College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China
[2]Changchun Institute of Optics, Fine Mechanics, and Physics, Chinese Academy of Sciences, Changchun 130033, China
[3]College of Communication Engineering, Jilin University, Changchun 130012, China

Corresponding author: Min Guo (guomin7852@163.com)

This work was supported in part by the 13th Five-Year National Key Research and Development Plan under Grant 2017YFB0404800, in part by the National Natural Science Foundation of China (NSFC) under Grant 61631009, in part by the Open Project of the National Engineering Laboratory under Grant 2018NELKFKT03, in part by the Science and Technology Project of Changchun under Grant 17DY007, and in part by the project named Technology Development of 3D Scanning Based on Structured Light.

ABSTRACT Limited available methods for the generation of an elemental image array (EIA) may be the most severe bottleneck for the promotion and application of integral imaging. To obtain rich information of three dimensional (3D) real scene at a relatively low cost, and solve the pseudoscopic problem in real display mode, we propose an EIA generation method based on convergent camera array through establishing the optical mapping model between the proposed method and the classical computer virtual generation method. We expound the basic principle of solving the pseudoscopic problem by changing the relative position of object and center of the convergent camera array in this paper. The experimental results indicate that the generated EIA can be displayed without the pseudoscopic problem in real display mode, relative position relationships of the objects in 3D real scene can also be accurately restored. Both preliminary optical experiments and the optical analysis are conducted to prove the feasibility and validity of the proposed method.

INDEX TERMS Elemental image array, convergent camera array, pseudoscopic problem, real display mode, optical mapping model.

I. INTRODUCTION

Integral imaging, which is modified from integral photography invited by Lippmann [1], offers a passive and relatively inexpensive way to capture 3D information and visualize it optically or computationally, it has attracted much attention due to its various advantages, for it can provide quasi-continuous, full-parallax and full-color 3D images with motion parallax but without coherent illumination, and it does not need any eye-tracking or visual aids [2]–[5].

A complete integral imaging should consist of pickup, storage, transmission, and display, and may also include the processes of image compression and decompression [3]–[5]. This paper focus on the main two parts: pickup and display, as shown in Fig. 1. In the pickup part shown in Fig. 1(a), lens array which generally refers to camera array for still scene and camcorder array for movable scene is used to sample the 3D scene. Information of the 3D scene such as the outline, color, brightness and depth information of the object are recorded as an EIA by recording media such as film, printing, and electronic photosensitive device. Each elemental image (EI) contains information of the object observed from different positions. In the display part shown in Fig. 1(b), EIA is displayed on a display device, such as a liquid-crystal display, rays emitted from the EIA are modulated by the lens array and propagated to reconstruct a 3D scene in the real world. The corresponding parameters of this set of lens array should be consistent with that in the pickup process to avoid depth distortion during displaying the 3D image [6].

As the most basic display content, EIA is critical for the development of integral imaging. The first approach for the capture of the spatial-angular information of rays emitted by a 3D scene was reported by Lippmann under the name of Integral Photography (IP) [1]. Since his original proposal, other alternative but equivalent methods have been reported in order to perform the same kind of capture as Lippmann [2], [3]. So far, there have mainly been two kinds of methods to generate the display content for integral imaging: computer

The associate editor coordinating the review of this manuscript and approving it for publication was Yongqiang Zhao.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

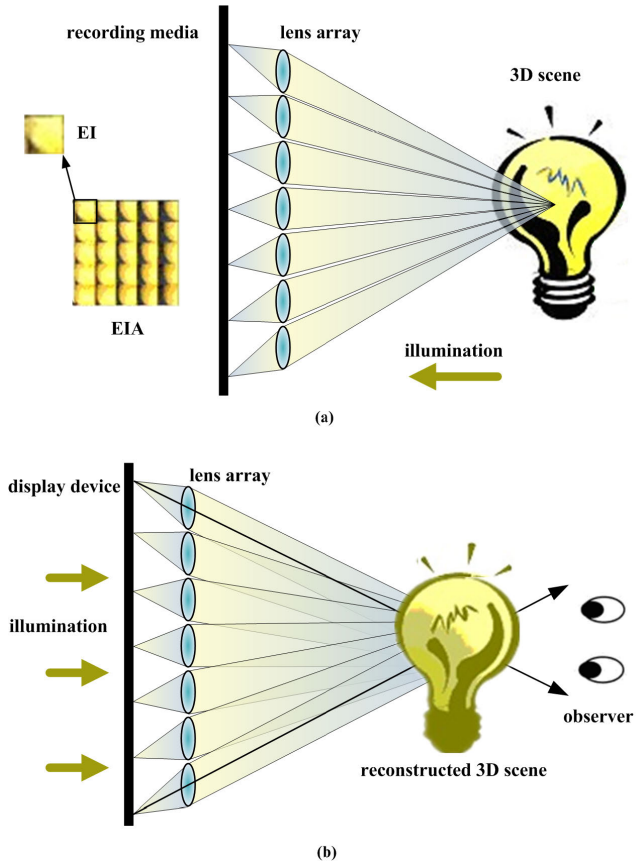

FIGURE 1. Two parts of integral imaging: (a) the pickup part, and (b) the display part.

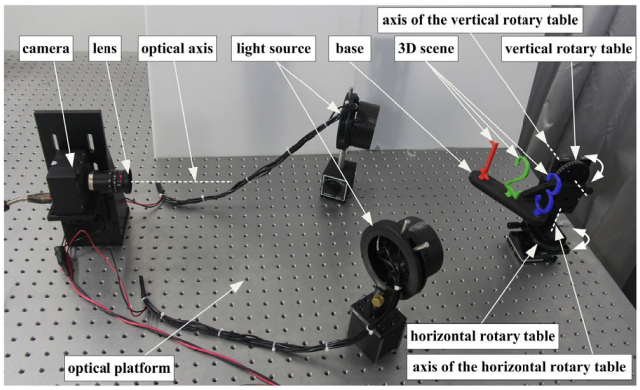

FIGURE 2. Convergent viewpoint image pickup system.

virtual synthesis and real scene acquisition [3]–[5]. Computer virtual synthesis is used to generate an EIA of virtual scenes by using some computer graphical software packages, such as 3D Studio Max and Maya [7]–[9]. However, real scene acquisition is preferable, because it does not need the time-consuming modeling process [10], [11]. As the most direct and simplest way of real scene acquisition, lens-based pickup scheme can be used to capture real scene directly with a lens array and a single recording medium, various approaches are proposed to improve the quality of EIA [3], [4], [11], [12], but a trade-off relationship exists between the display resolution and the number of EIs for a given sensor device. Using a camera or sensor array can solve the problem effectively, and each camera or sensor records an EI of the object from a different location. However, the scale of the array increases along with the increase of the resolution and the scale of real scene, resulting in higher costs and more workload [13]–[15].

A sub-image array (SIA) is composed of a series of viewpoint images, and it can be transformed into a corresponding EIA according to the optical mapping relationship [16]. Based on this, in this paper, we propose a generation method of an EIA for integral imaging using a convergent viewpoint pickup system and the optical mapping model between the proposed method and the classical computer virtual synthesis.

The structure of this paper is as follows: the convergent viewpoint pickup system is described in Section II, theoretical principle and the generation process are discussed in Section III, section IV covers solution to the pseudoscopic problem in the real display mode, experimental results and contrastive analysis are presented in Section V, and conclusions are drawn in Section VI.

II. CONVERGENT VIEWPOINT IMAGE PICKUP SYSTEM

In an ideal spherical camera array, all cameras are on the same sphere, all cameras' optical axis are pointing to the center of sphere, and the object is placed in the public field near the center to make sure it can be completely captured by each camera. We build an equivalent pickup system to the spherical camera array, as shown in Fig. 2, it mainly includes the camera, lens, standard light source, and the vertical and horizontal rotary tables, all the equipment are fixed on the optical platform by optical connectors. The color camera we used, DH-SV401FC, is produced by 'Da-Heng Image' with 1394 interface, image resolution 780 pixels×582 pixels, pixel size 8.3 um×8.3 um, images are transmitted from the camera to the host computer through the 1394 image acquisition card, the lens modeled LM8JC10M is 8.5mm prime lens, and it is produced by KOWA.

The vertical rotary table and the horizontal rotary table make up the two-dimensional (2D) rotary table, the former can make the object goes around its axis (the yellow dotted line) which is parallel to the optical platform, and the latter can make the object goes around its axis (the red dotted line) which is perpendiculars to the optical platform. The axes of those two rotary tables intersect on the axis of the camera, the intersection point is the center of the 2D rotary table, and it is approximately 800 mm away from the camera lens. There will be relative movement between the camera and the object when we rotate the horizontal or the vertical rotary table. During the movement, the optical axis of the camera (the white dotted line) always points to the center of the 2D rotary table, and distance between the camera and the center remains unchanged. Therefore, the center of the 2D rotary table is the equivalent of the center of the spherical camera

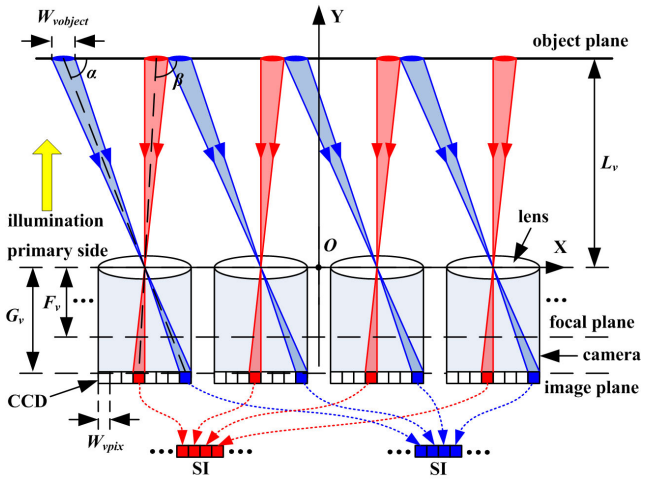

FIGURE 3. Schematic diagram of computer virtual pickup method.

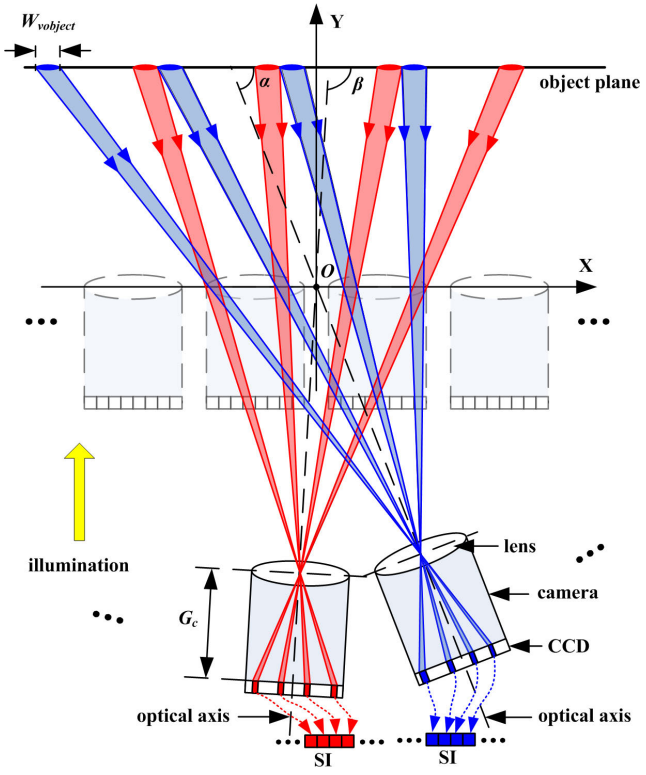

FIGURE 4. Schematic diagram of the convergent camera array pickup method.

array. We can realize the equivalent collection functionality of the spherical camera array by adjusting the 2D rotary table to a specific set of rotating angles.

III. THEORETICAL PRINCIPLE AND GENERATING PROCESS OF THE PROPOSED METHOD

A. THE OPTICAL MAPPING MODEL OF COMPUTER VIRTUAL ACQUISITION METHOD AND CONVERGENT CAMERA ARRAY ACQUISITION METHOD

To provide theoretical basis for the design and implementation of the convergent viewpoint pickup system and prove the validity of the proposed method, we establish the optical mapping model between the convergent camera array acquisition method and the computer virtual pickup method which is a classical generation method of EIA.

In the computer virtual pickup method, as shown in Fig. 3, we take four cameras for example, the axes of all cameras are parallel to each other, and all the cameras' centers are located in the same line, every camera can collect one EI. We set point A as the center of the virtual camera array, the red area of the object plane is imaged as the 4-*th* pixel on the corresponding CCD, and the blue area of the object plane is imaged as the 8-*th* pixel on the corresponding CCD. The angle between the object plane and the line, which is defined by the center of the blue area and the center of the corresponding pixel area, is α, as shown in Fig. 3, and the angle between the object plane and the line, which is defined by the center of the red area and the center of the corresponding pixel area, is β. Then all the red areas correspond to one sub-image (SI), and all the blue areas correspond to another SI. Since there are 8 pixels on the CCD of every camera, the EIA captured by the virtual camera array shown if Fig. 3 can be transformed to 8 SIs according to the optical mapping relationship between SIA and EIA [16].

To obtain the same 8 SIs as those in Fig. 3 using the proposed method, we need a convergent camera array composed by 8 cameras. For convenience, the image collected by the convergent camera is called convergent viewpoint image, and the pixels that make up the SIs are called valid pixels. Take 2 cameras for example as shown in Fig. 4, the intersection A of the two cameras' optical axes, namely the center the convergent camera array, is equivalent to point A in Fig. 3. There is a one-to-one correspondent relationship between the convergent viewpoint image and the SI, as you can see in Fig. 4, the angle between the object plane and the convergent camera's optical axis is equal to the angle between the object plane and the imaging light of pixel from the corresponding SI, that is the α in Fig. 3 is the same as that in Fig. 4, and the β in Fig. 3 is also the same as that in Fig. 4. We can see that the target pixels are scattered in the convergent viewpoint image, as shown in Fig. 4, and the target pixel's width is different from each other. To extract the corresponding SI from the convergent viewpoint image, we should first accurately calculate the relative position of the target pixel in the corresponding convergent viewpoint image. There are two specific parameters: the center coordinates of the convergent viewpoint image and the boundary point coordinates of the target pixel, and the detailed calculation processes are as follows.

1) CALCULATE THE CENTER COORDINATE OF THE CONVERGENT VIEWPOINT IMAGE

First, we set up coordinate systems with point A as the origin in both Fig. 3 and Fig. 4, X-axes are all parallel to the object plane and Y-axes are all perpendicular to the X-axes. Suppose

each camera's CCD includes N_{vccd} pixels, the size of each pixel is W_{vpix}, then N_{vccd} cameras should be included in the convergent camera array, then the angle between the object plane and the *k-th* camera's optical axis can be expressed as

$$\theta_k = \arctan\left\{G_v/[(k-0.5)\times W_{vpix} - N_{vccd}\times W_{vpix}/2]\right\}, \tag{1}$$

where, arctan(·) represents the inverse tangent for the value in parentheses.

Suppose the distance from the center of camera lens to point A in Fig. 4 is R, then the coordinate of the lens center of the *k-th* convergent camera, timely (XCC_k, YCC_k), can be formulated as

$$\begin{cases} XCC_k = R\times\cos(\theta_k) \\ YCC_k = -|R\times\sin(\theta_k)| \end{cases}, \tag{2}$$

where, sin(·) represents the sine for the value in parentheses. Use (XCM_k, YCM_k) to describe the coordinate of the intersection point of the *k-th* convergent camera's optical axis and CCD, namely the center of the convergent viewpoint image. Then the intersection point and the lens center should satisfy the following relationship according to the optical imaging relationship:

$$\frac{XCC_k}{YCC_k} = \frac{XCM_k}{YCM_k}, \tag{3}$$

$$(XCM_k - XCC_k)^2 + (YCM_k - YCC_k)^2 = G_c^2, \tag{4}$$

where, G_c is the imaging distance of the convergent camera as shown in Fig. 4.

2) CALCULATE THE COORDINATES OF THE TARGET PIXELS

Suppose that there are N_{vc} virtual cameras in the virtual camera array, as shown in Fig. 3, the distance between the optical axes of two adjacent virtual cameras is D_{vc}, then the Y-coordinate of the lens center of the *m-th* virtual camera in the virtual camera array is 0, and the X-coordinate XVC_m can be expressed as

$$XVC_m = D_{vc}\times(m-1) - (N_{vc}-1)\times D_{vc}/2. \tag{5}$$

The Y-coordinate of the imaging region center of the *k-th* pixel on the CCD of the *m-th* virtual camera array is $-G_v$, and the X-coordinate is

$$XVP_{mk} = XVC_m + W_{vpix}\times(k-1) - (N_{vccd}-1)\times W_{vpix}/2. \tag{6}$$

The width of each red or blue area in Fig. 4 is equal to $W_{vobject}$, which is the spatial resolution of virtual camera at the object plane, and it can be calculated according to Gaussian imaging formula

$$W_{vobject} = W_{vpix}\times L_v/G_v. \tag{7}$$

The imaging area of the *k-th* pixel on the CCD of the *m-th* virtual camera corresponds to the region on the object plane which is centered around (XVO_{mk}, L_v), and the width is $W_{vobject}$, XVO_{mk} can be calculated according to the geometric relations in Fig. 3 and Fig. 4

$$XVO_{mk} = XVC_m - [(XVP_{mk} - XVC_m)\times L_v]/G_v. \tag{8}$$

The specific expressions of XCM_k and YCM_k can be obtained from (7) and (8). Set the coordinates of the left boundary point of the imaging region, which is corresponding to the *m-th* pixel on the *k-th* SI, on the *k-th* convergent camera's CCD in Fig. 4 as (XCL_{mk}, YCL_{mk}), and the right boundary point (XCR_{mk}, YCR_{mk}). Then the following expressions can be obtained with optical imaging relationship

$$\frac{(XVO_{mk} + W_{vobject}/2) - XCC_k}{L_v - YCC_k} = \frac{XCC_k - XCL_{mk}}{YCC_k - YCL_{mk}}, \tag{9}$$

$$\frac{(XVO_{mk} - W_{vobject}/2) - XCC_k}{L_v - YCC_k} = \frac{XCC_k - XCR_{mk}}{YCC_k - YCR_{mk}}, \tag{10}$$

$$\begin{aligned}&(XCL_{mk} - XCM_k)(XCC_k - XCM_k)\\ &+(YCL_{mk} - YCM_k)(YCC_k - YCM_k) = 0,\end{aligned} \tag{11}$$

$$\begin{aligned}&(XCR_{mk} - XCM_k)(XCC_k - XCM_k)\\ &+(YCR_{mk} - YCM_k)(YCC_k - YCM_k) = 0.\end{aligned} \tag{12}$$

The specific values of XCL_{mk}, YCL_{mk}, XCR_{mk}, and YCR_{mk} can be calculated from (9)-(12). Coordinate of the intersection of the *k-th* convergent camera's optical axis and the CCD, timely (XCM_k, YCM_k), can also be calculated. With this, we establish the mapping model between SIA and the convergent viewpoint image array (CVIA) by accurately calculating the position coordinates of the valid pixels that constitute one SI in the corresponding convergent viewpoint image. Given that SIA can be transformed into EIA which can be obtained directly using computer virtual acquisition method, and the CVIA can be obtained with the proposed convergent camera array acquisition method, therefore, the optical mapping model between computer virtual acquisition method and the convergent camera array acquisition method is finally established in this paper.

B. EIA GENERATION WITH THE PROPOSED METHOD

The process of generating EIA using the optical mapping model described above mainly includes the following four steps as show in Fig. 5. Firstly calculate the hardware parameters of the convergent camera array according to the hardware parameters of the display platform and build the convergent camera array accordingly. Secondly collect the CVIA of the target 3D real scene, calibrate the camera for the internal and external parameters using the camera calibration method based on Zhang Zhengyou calibration method [17] and correct the distortion of every convergent viewpoint image with the calibration results. Thirdly build the optical mapping model between CVIA and SIA through calculating the coordinates of valid pixels in CVIA and extracting the valid pixels from the distortion-corrected convergent viewpoint images, and generate SIs with the extracted valid pixels according to

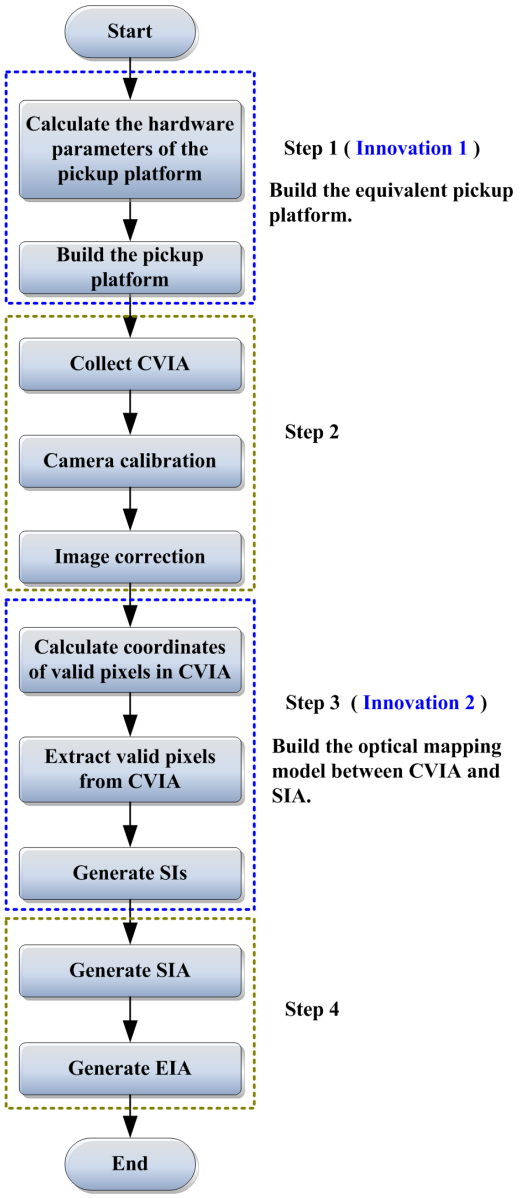

FIGURE 5. Schematic diagram of the proposed method.

the built optical mapping model. Finally, transform SIA into the corresponding EIA.

IV. SOLUTION TO THE PSEUDOSCOPIC PROBLEM

In 1933, Ives from Bell Telephone Laboratories investigated the pseudoscopic problem [2], which is one of the original drawbacks of integral imaging to reconstruct the integrated 3D image in real display mode, and it is the phenomenon that the location of the reconstructed image is inverted back to front. The real display mode and the virtual display mode are the common display modes in integral imaging, and Fig. 6 shows the cause and formation of the pseudoscopic phenomenon. As shown in Fig. 6(a), the direction of pickup is the same as that of display, so the pseudoscopic problem does not exist in the virtual display mode. But for the real display

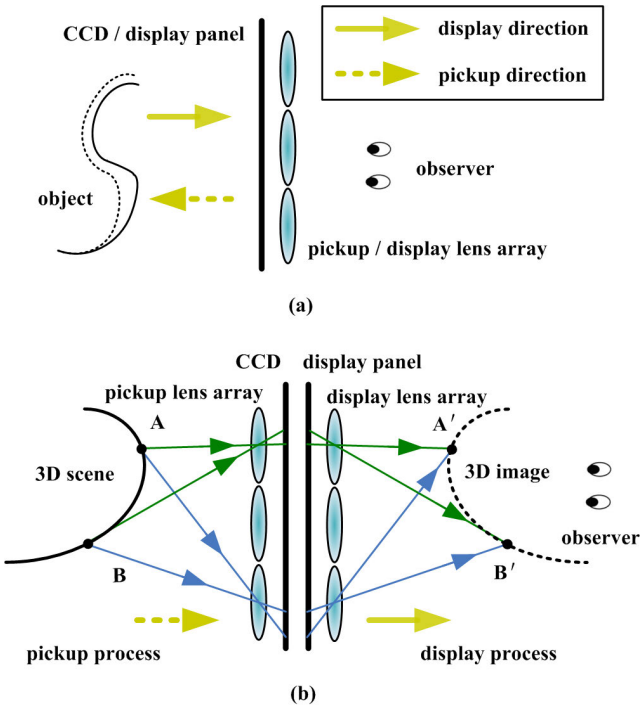

FIGURE 6. Directions of the pickup and the display processes for (a) virtual display mode, and (b) real display mode.

mode shown in Fig. 6(b), the 3D image, whose light field information is the same as that of the 3D scene, is recovered according to the reversible principle of the optical path. Point A is in front of point B in the 3D scene, points A′ and B′ are the reconstructed images of points A and B respectively, but point A′ is behind point B′ in the reconstructed 3D image. The cause is that the direction of pickup is opposite to that of display, and the image of the object nearer to the pickup device is located farther from the observer in the real display mode [5], [18]. Given the fact that the directions of pickup and display are the same in the virtual display mode, and the pseudoscopic problem vanishes, common approaches to the problem is to rotate each elemental image 180°center-symmetrically, then EIA is converted into that for the virtual display mode [11], but the problem is merely avoided but not solved. Dual-lens-array can also be used to generate the pseudoscopic-free EIA through two shooting processes, but it puts forward a higher request to the recent manufacture technology of lens array and the operational accuracy [4], [2].

Actually, 3D image in the real display mode is closer to the observer, which can bring greater visual impact and influences in 3D display process. Therefore, it is meaningful to generate the pseudoscopic-free EIA of a 3D real scene for the real display mode using the optical pickup method.

According to the above analysis, if the pickup direction in Fig. 6(b) goes in the opposite direction, which means the shooting direction is the same as the display direction as shown in Fig. 7, then the generated EIA should be displayed without pseudoscopic problem in the real display mode.

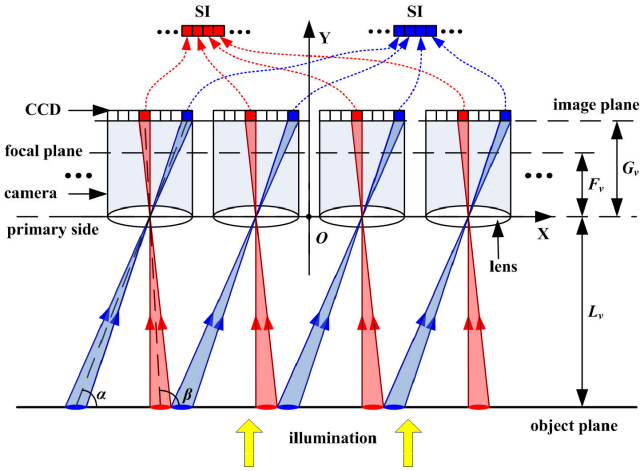

FIGURE 7. Schematic diagram of the imaging light path of the camera array when the shooting direction is the same as the viewing direction.

However, the imaging optical path depicted in Fig. 7 is not tenable in fact, because the reflected light from the object cannot continue to travel in the direction of the light source when the illumination falls on an opaque object on the object plane. Therefore, it is impossible for the camera behind the object to collect images of the front of the object, and EIA corresponding to the SIs in Fig. 7 can not be directly obtained by traditional optical acquisition method. However, with the proposed method, the SIs of the front face of the object can be obtained from the convergent cameras, and then the corresponding EIA can be obtained according to the optical mapping relationship between SIA and EIA. The basic principle of the proposed method is shown in Fig. 8.

The optical mapping mode shown in Fig. 8 can be obtained using the same derivation process as Section III (A). It is important to note that the longitudinal directions in Fig. 3 and Fig. 4 are opposite to the imaging light direction in Fig. 3, but the longitudinal directions in Fig. 7 and Fig. 8 are the same as the imaging light direction in Fig. 7. Since the origin of the coordinate system is the intersection point of all the convergent cameras' optical axes, timely the point O, the following conclusions can be drawn through analyzing the corresponding relationship between the two coordinate systems shown in Fig. 7 and Fig. 8: (1) when the point O located between the object plane and the convergent camera array, the generated EIA can be displayed normally in the virtual display mode; (2) when the object is located between point O and the convergent camera array, the generated EIA can be displayed without pseudoscopic problem in the real display mode. Situation (1) refers to the virtual display mode which is used in most of the existing EIA generation methods for 3D real scene. And we focus on situation (2) in this paper.

V. EXPERIMENTAL RESULTS AND ANALYSIS

For the pickup platform shown in Fig. 2, three entity numbers '1', '2', and '3', whose planes are parallel to each other and the interval is 30mm, are used as the 3D real scene and fixed

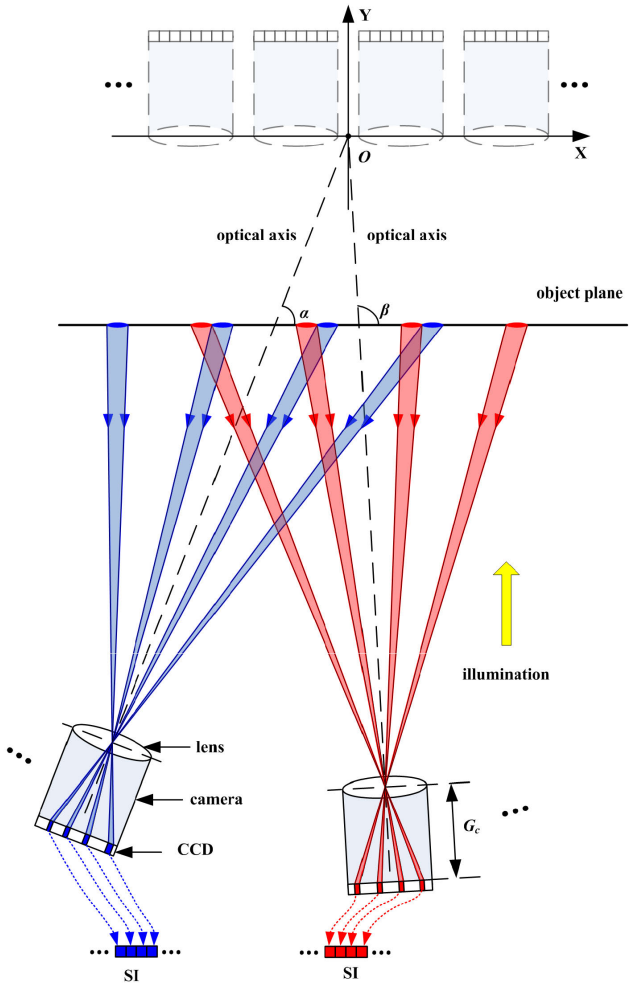

FIGURE 8. Basic principle for solving the pseudoscopic problem.

vertically on the black base. The black base is fixed on the vertical rotary table, whose axis is parallel to the plane of number '3' and the interval is 80mm. The display platform used in the experiment is shown in Fig. 9, in which there are 5 (H)×8 (V) agglutination lens, and the effective focal length is 30.07 mm, the clear aperture is 15 mm. The distances between centers of two adjacent agglutination lens in the horizontal and vertical directions are both 18.6 mm. The agglutination lens array is 40 mm away from the liquid crystal flat panel display whose resolution is 1366 pixels×768 pixels and the pixel size is 0.3 mm×0.3 mm. As the displayed real images can't be filmed directly by a camera, we use the sulfate paper screen, as shown in Fig. 9, to receive the real image displayed to present us the display effect intuitively.

The 15 values of the rotation angles are calculated from (1), as shown in Table 1. The signs of the angle values are defined in line with the following rule: when we are facing the axis of the rotary table, the clockwise rotation of the rotary table is positive, and the counterclockwise is negative. In the pickup process, 225 convergent viewpoint images are collected according to the corresponding relationship shown

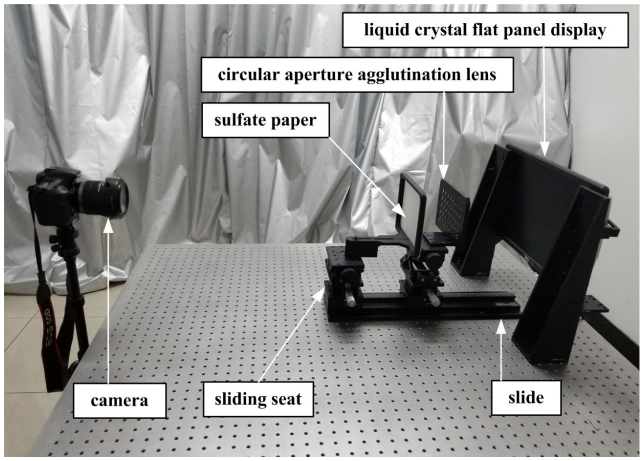

FIGURE 9. Display platform for the proposed method.

TABLE 1. Values of the rotation angles.

Symbol 1	Symbol 2	Values
θ_1	$-\theta_{15}$	-11.87°
θ_2	$-\theta_{14}$	-10.21°
θ_3	$-\theta_{13}$	-8.54°
θ_4	$-\theta_{12}$	-6.85°
θ_5	$-\theta_{11}$	-5.15°
θ_6	$-\theta_{10}$	-3.44°
θ_7	$-\theta_9$	-1.72°
θ_8		0°

TABLE 2. Specific correspondence of rotation angle values of 2D turntable and the convergent viewpoint image numbers.

	θ_1	θ_2	θ_3	θ_4	⋯	θ_{12}	θ_{13}	θ_{14}	θ_{15}
θ_1	15.1	15.2	15.3	15.4	⋯	15.12	15.13	15.14	15.15
θ_2	14.1	14.2	14.3	14.4	⋯	14.12	14.13	14.14	14.15
θ_3	13.1	13.2	13.3	13.4	⋯	13.12	13.13	13.14	13.15
θ_4	12.1	12.2	12.3	12.4	⋯	12.12	12.13	12.14	12.15
⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮
θ_{12}	4.1	4.2	4.3	4.4	⋯	4.12	4.13	4.14	4.15
θ_{13}	3.1	3.2	3.3	3.4	⋯	3.12	3.13	3.14	3.15
θ_{14}	2.1	2.2	2.3	2.4	⋯	2.12	2.13	2.14	2.15
θ_{15}	1.1	1.2	1.3	1.4	⋯	1.12	1.13	1.14	1.15

in Table 2, in which the '15.1' represents the image located in row 15 and column 1 of the convergent viewpoint image array.

The distortion-corrected convergent viewpoint images numbered 1.1, 1.15, 8.8, 15.1, and 15.15 are shown in Fig. 10, from which we can also see that the perspectives of the obtained image information of 3D real scene with the pickup

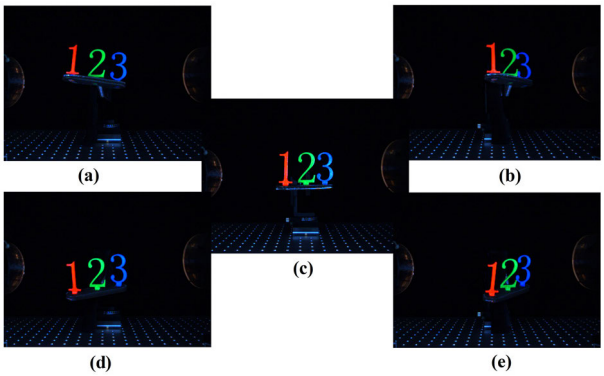

FIGURE 10. Distortion-corrected convergent viewpoint images: (a), (b), (c), (d), and (e) for images numbered 1.1, 1.15, 8.8, 15.1, and 15.15.

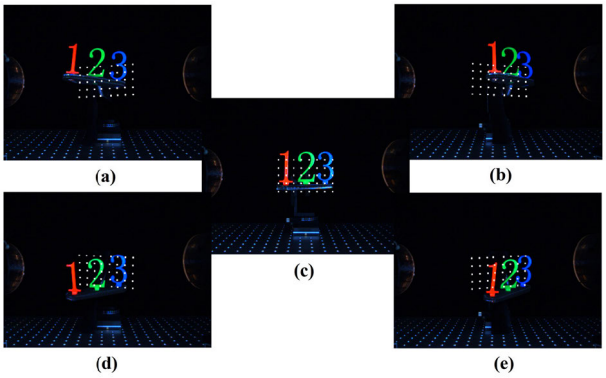

FIGURE 11. Distribution of SI pixels in the convergent viewpoint images: (a), (b), (c), (d) and (e) are for (a), (b), (c), (d), and (e) in Fig. 10.

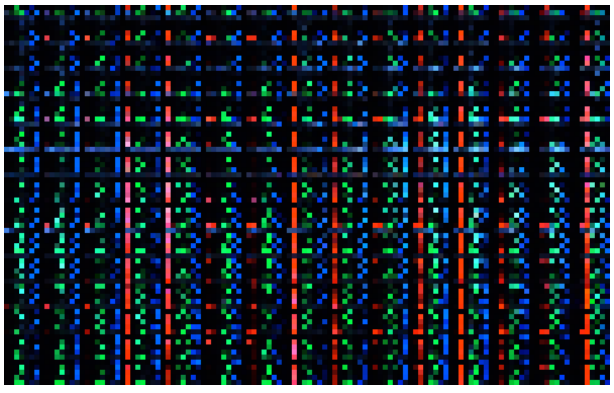

FIGURE 12. SIA of the proposed method.

system we set up are equal to those of a convergent camera array. Distribution of valid pixels in the five convergent viewpoint images in Fig. 10 is shown in Fig. 11, and the white rectangle represents the extraction area of a pixel.

SIA generated with the extracted valid pixels is as shown in Fig. 12, which includes 15 × 15 SIs with the resolution of 5 pixels×8 pixels. The corresponding EIA shown in Fig.13 includes 5 × 8 EIs, and it is obtained according to the optical mapping relationship between SIA and EIA.

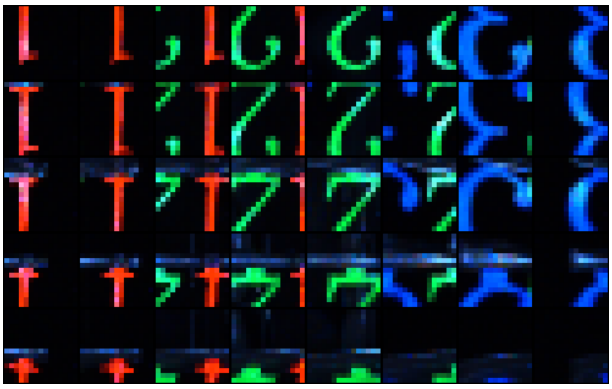

FIGURE 13. EIA of the proposed method.

The generated EIA is displayed on the display platform shown in Fig. 9. In order to fully demonstrate the display effect of the real image, we use the sulfate paper screen to receive the imaging light in space, and then the camera is used to shoot the image on the sulfate paper. So the image collected by the camera is the scene that the human eye sees when focusing. During the process of moving the sulfate paper curtain, images on the sulfate paper captured by the camera when the distance between the sulfate paper and the agglutination lens array is separately 140 mm, 110 mm, and 80 mm are shown in Fig. 14, from which we can see that: 1) the image of the number '1' is clear, and the images of '2' and '3' are blurred in Fig. 14(a); 2) the image of the number '2' is clear, and the images of '1' and '3' are blurred in Fig. 14(b); 3) the image of the number '3' is clear, and the images of '1' and '2' are blurred in Fig. 14(c). It is obviously that the intervals of the three numbers' optimum imaging planes are 30 mm, and the optimum imaging plane of number '3' is 80 mm away from the agglutination lens array, all those data are all completely consistent with the actual pickup scene. It can also be seen that when the object is located between the camera and the center of the 2D rotary table, the proposed method can be used to generate EIA with the correct depth for the real display mode, the positions of the three optimal imaging planes quantitatively proves the effectiveness of the proposed method.

As to the EIA generation method of large size 3D real scene, we previously proposed an EIA generation method using a parallel camera array in a paper published in the journal of Applied Optics [6], the corresponding discrete viewpoint image array acquisition system, which is based on monocular camera and slide tracks, as shown in Fig. 15. The acquisition system mainly includes camera, lens, two light sources, shooting object, vertical slide track and horizontal slide track. The shooting object is constituted of two entity letters, the red letter 'B' is placed in front of the green letter 'H', and the red letter 'B' blocks the green letter 'H'. The camera is mounted on the slide seat of the vertical slide rail, which makes the camera move linearly in the direction perpendicular to the plane of the optical

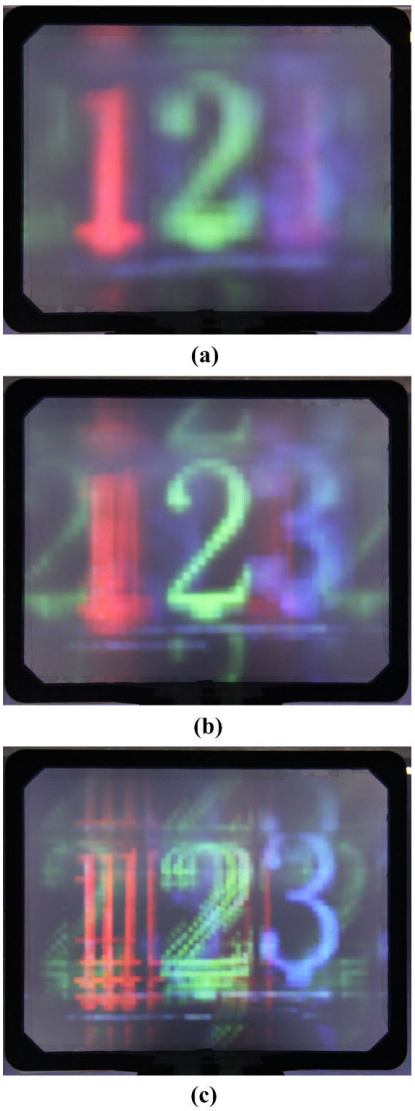

FIGURE 14. The captured images when the sulfate paper is (a) 140 mm, (b) 110 mm, and (c) 80 mm away from the agglutination lens array.

platform, and the horizontal slide make the camera move in the direction parallel to the optical platform. The coordinates of the camera on the horizontal and vertical slide tracks can be obtained quantitatively through the rulers on the slide tracks. Vertical slide rail, horizontal slide rail and the two sliding seats used for connection make up the 2D manual translation stage, which can make the camera move freely on the plane which is vertical to the optics platform. With this 2D manual translation stage, we can obtain the equivalent images of static scene to those of the discrete camera array by respectively moving the camera to multiple specific locations of the plane and capturing the images. And then we use the interception window to intercept the valid pixels from the obtained discrete viewpoint images to generate the corresponding SIA and EIA as shown in Fig. 16, and the display effect is as shown in Fig. 17. It can be seen that the letter 'B' blocks obviously the letter 'H', the left vertical lines of the green

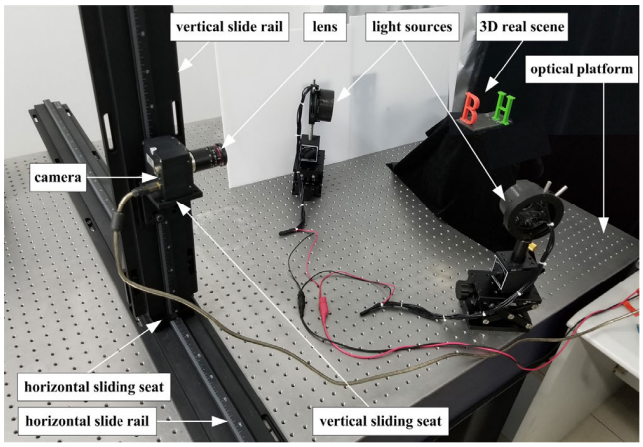

FIGURE 15. Pickup platform of the method in [6].

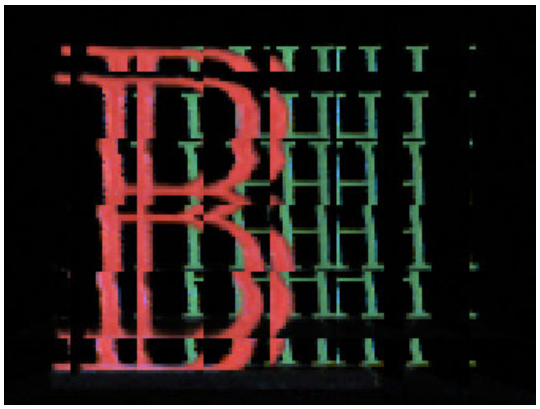

FIGURE 16. EIA generated by the method in [6].

letter 'H' can be seen through the two holes of the red letter 'B' when we view the displayed image from the left side, as shown in Fig. 17(a). Correspondingly, the letter 'B' blocks the letter 'H' slightly when we view the displayed image from the right side, as shown in Fig. 17(b). The above phenomena fully meet the visual effects of human eyes in real world. The experimental results show that the display effect of the generated EIA can truly restore the 3D information of the real scene.

As mentioned above, we have proposed two EIA generation methods: the parallel camera array method and the convergent camera array method. Although we name them as the parallel and convergent camera array, we actually use only one camera in the whole pickup process, and the two methods can both generate EIA of large-size 3D real scene with low workload cost and correct depth information when displayed. Furthermore, they can both realize real-time generation and storage of EIA, improving the generation efficiency of the EIA, especially the video sequence, greatly compared with other traditional camera array systems.

In terms of similarities of the two EIA generation methods we proposed, there are mainly the following three aspects: (1) Both methods belong to the acquisition method combining optical acquisition and computer virtual synthesis, and they both belong to the EIA generation technology of real 3D scene; (2) as to the implementation steps, they both firstly

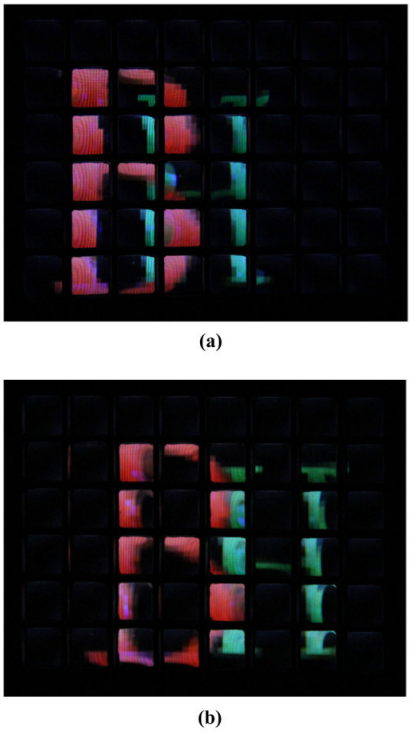

FIGURE 17. Display effect of EIA in Fig. 16: (a) left view and (b) right view.

build an optical mapping model, then extract the effective pixels, and finally generate EIA; (3) The optical mapping relationship between EIA and SIA is effectively used to significantly reduce the shooting cost and workload.

However, there are also differences between the two methods: (1) the previously proposed method is suitable for a parallel camera array while this method for a convergent camera array; (2) The complexity of this method is slightly higher than that of the former method. In the former method, since the size of the valid pixel and the space between the valid pixels in the discrete viewpoint images are the same, we can make the overall calculation, and put forward the concept of 'interception window', and then the size and location of the valid pixels in the discrete viewpoint images can be represented by the follow three parameters: the size, the initial position and the moving step parameters of the interception window. However, in the proposed algorithm in this paper, we calculate respectively the size and location of the valid pixels in the convergent viewpoint images according to the different positions of the shooting camera. As far as the computing power of the existing computers is concerned, these two methods can both realize real-time computing without affecting the generation of video sequences; (3) As we all know, the convergent camera array has larger public view than the parallel camera array in the process of binocular stereo collection when the camera parameter, the binocular baseline distance and the shooting distance are all fixed. Therefore, this method has a larger collection field, and can be used to collect EIA of a larger 3D scene. Meanwhile the light field information of real scene recorded by the EIA generated

by this method is more abundant, because the convergent camera array can provide more different angles, from which we can collect more optical information of the 3D real scene, also this kind of structure can solve the occlusion problem to some extent.

VI. CONCLUSION

In this paper, we propose an EIA generation method using a convergent camera array through the establishment of the optical mapping model between the convergent viewpoint image and the elemental image, build the optical acquisition model based on SI according to the mapping relationship between SI and EI, replace EI whose number is larger and resolution lower with SI whose number is smaller and resolution higher without changing the display resolution, significantly reducing the computational complexity and improving the generation efficiency of EIA. The proposed method maximizes the 3D scene optical information at a lower pickup cost compared with the traditional camera array acquisition method, and this method can generate EIA of a larger scale 3D scene without the complex modeling process compared with the computer virtual acquisition method. Both theoretical analysis and experimental results show that the obvious advantage of the proposed method is that the generated EIA can be displayed in both real display mode and virtual display mode without the pseudoscopic problem only by controlling the relative position relationship between the convergent camera array and the 3D object, which improves the richness of 3D display content for integral imaging. Actually, the collected images are slightly distorted as a result of the perspective projection, but this distortion is imperceptible to human eye, because the distance between the target 3D scene and the camera is much greater than the depth of field. In future work, we will explore the elimination of this kind of distortion to improve the accuracy of the collection method.

REFERENCES

[1] M. G. Lippmann, "La photographic integrale," *Comtes Rendus, Academie Des Sci.*, vol. 146, pp. 446–451, Sep. 1908.
[2] H. E. Ives, "Optical properties of a lippmann lenticulated sheet," *J. Opt. Soc. Amer.*, vol. 21, no. 3, pp. 171–176, Mar. 1931.
[3] Q. Ma, L. Cao, Z. He, and S. Zhang, "Progress of three-dimensional light-field display [invited]," *Chin. Opt. Lett.*, vol. 17, no. 11, 2019, Art. no. 111001.
[4] M. Martinez-Corral, A. Dorado, J. C. Barreiro, G. Saavedra, and B. Javidi, "Recent advances in the capture and display of macroscopic and microscopic 3-D scenes by integral imaging," *Proc. IEEE*, vol. 105, no. 5, pp. 825–836, May 2017.
[5] M. Martínez-Corral and B. Javidi, "Fundamentals of 3D imaging and displays: A tutorial on integral imaging, light-field, and plenoptic systems," *Adv. Opt. Photon.*, vol. 10, no. 3, pp. 512–566, Sep. 2018.
[6] M. Guo, Y. Si, Y. Lyu, S. Wang, and F. Jin, "Elemental image array generation based on discrete viewpoint pickup and window interception in integral imaging," *Appl. Opt.*, vol. 54, no. 4, pp. 876–884, Feb. 2015.
[7] A. Chutjian and R. J. Collier, "Recording and reconstructing three-dimensional images of computer-generated subjects by Lippmann integral photography," *Appl. Opt.*, vol. 7, no. 1, pp. 99–103, Jan. 1968.
[8] S.-L. Li, Q.-H. Wang, Z.-L. Xiong, H. Deng, and C.-C. Ji, "Multiple orthographic frustum combing for real-time computer-generated integral imaging system," *J. Display Technol.*, vol. 10, no. 8, pp. 704–709, Aug. 2014.
[9] K. S. Park, S.-W. Min, and Y. Cho, "Viewpoint vector rendering for efficient elemental image generation," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 1, pp. 233–241, Jan. 2007.
[10] F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography," *Appl. Opt.*, vol. 36, no. 7, pp. 1598–1603, Mar. 1997.
[11] J. Arai, F. Okano, H. Hoshino, and I. Yuyama, "Gradient-index lens-array method based on real-time integral photography for three-dimensional images," *Appl. Opt.*, vol. 37, no. 11, pp. 2034–2045, Apr. 1998.
[12] J. Wang, X. Xiao, G. Yao, A. Stern, and B. Javidi, "Synthetic aperture integral imaging display with moving array lenslet technique," *J. Display Technol.*, vol. 11, no. 10, pp. 827–833, Oct. 2015.
[13] H. Ren, Q.-H. Wang, Y. Xing, M. Zhao, L. Luo, and H. Deng, "Super-multiview integral imaging scheme based on sparse camera array and CNN super-resolution," *Appl. Opt.*, vol. 58, no. 5, pp. A190–A196, Feb. 2019.
[14] C.-G. Luo, Q.-H. Wang, H. Deng, and Y. Liu, "Extended depth-of-field in integral-imaging pickup process based on amplitude-modulated sensor arrays," *Opt. Eng.*, vol. 54, no. 7, Jul. 2015, Art. no. 073108.
[15] S.-H. Hong, J.-S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Express*, vol. 12, no. 3, pp. 483–491, Feb. 2004.
[16] J.-H. Park, J. Kim, and B. Lee, "Three-dimensional optical correlator using a sub-image array," *Opt. Express*, vol. 13, no. 13, pp. 5116–5126, Jun. 2005.
[17] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Dec. 2000.
[18] C. B. Burckhardt, R. J. Collier, and E. T. Doherty, "Formation and inversion of pseudoscopic images," *Appl. Opt.*, vol. 7, no. 4, pp. 627–632, Apr. 1968.

MIN GUO received the B.E. degree in electronic information engineering from the Changchun University of Technology, in 2009, and the M.E. degree in signal and information processing and the Ph.D. degree in communication and information systems from Jilin University, in 2012 and 2018, respectively.

From 2012 to 2013, she was a Research Assistant with Changchun Shikai Technology Industry Limited Liability Company. Since 2019, she has been a Lecturer with the Department of Electronic Information Engineering, Changchun University of Technology. She is the author of eight articles and holds one invention. Her research interests include 3D image processing and machine vision.

YUANZHI LYU received the B.E. degree in optical information science and technology and the Ph.D. degree in communication and information systems from Jilin University, in 2009 and 2014, respectively.

From 2014 to 2019, he was a Research Assistant with the Changchun Institute of Optics, Fine Mechanics, and Physics, Chinese Academy of Sciences, where he has been an Associate Research Fellow, since 2019. He is the author of 17 articles and holds 13 patents of invention. His research interests include image processing, machine vision, and true-3D display.

SHIGANG WANG received the B.E. degree in automation technology from Dongbei University, in 1983, and the M.E. degree in communication and electronic systems and the Ph.D. degree in communication and information systems from Jilin University, in 1998 and 2001, respectively.

From 1985 to 1996, he was a Lecturer with the Changchun Institute of Posts and Telecommunications and an Associate Professor, from 1996 to 2001. Since 2001, he has been a Professor with Jilin University. He is the author of more than 100 articles and holds more than 40 patents of invention. His research interests include 3D image processing, recognition and understanding of digital image, and true 3D display.