photonics

MDPI

Article

A Sparse Capture Light-Field Coding Algorithm Based on Target Pixel Matching for a Multi-Projector-Type Light-Field Display System

Qingyu Meng [1], Haiyang Yu [2], Xiaoyu Jiang [2,*] and Xinzhu Sang [1,*]

[1] State Key Laboratory of Information Photonics and Optical Communication, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2] Army Academy of Armored Forces, Beijing 100072, China
* Correspondence: 13910593731@163.com (X.J.); xzsang@bupt.edu.cn (X.S.)

Abstract: The traditional light-field coding algorithm used in a multi-projector-type light-field display system requires sophisticated and complex three-dimensional modeling processes or parallax images obtained through dense capture. Here we propose an algorithm based on target pixel matching, which directly uses parallax images without a complex modeling process, and can achieve a more accurate light-field reconstruction effect under sparse capture conditions. For the lack of capture information caused by sparse capture, this algorithm compares the pixel similarity of the captured images of the same object point on different cameras to accurately determine the real capture information of the object point at different depths, which is recorded as the target pixel, and then the target pixel is encoded according to the lighting path to obtain the correct projector image array (PIA). By comparing the quality of PIAs generated by the traditional light-field coding algorithm and the display effect after loading the PIAs into the actual display system, we proved the effectiveness of the algorithm.

Keywords: light-field coding algorithm; sparse capture; multi-projector-type light-field display system; projector image array

Citation: Meng, Q.; Yu, H.; Jiang, X.; Sang, X. A Sparse Capture Light-Field Coding Algorithm Based on Target Pixel Matching for a Multi-Projector-Type Light-Field Display System. *Photonics* **2023**, *10*, 223. https://doi.org/10.3390/photonics10020223

Received: 26 December 2022
Revised: 13 February 2023
Accepted: 17 February 2023
Published: 19 February 2023

1. Introduction

Light-field display [1] is a new three-dimensional (3D) display technology that gives the observer an immersive stereoscopic visual experience by reconstructing the real light field, which includes integrated imaging technology [2–4], grating 3D display technology [5–7], holographic 3D display technology [8–10], and multi-projector-type 3D display technology [11–14]. The multi-projector-type 3D display technology has excellent development potential and advantages owing to its easily expanded space bandwidth product. This technology mainly consists of a projector array and a light-regulating device, which projects a large amount of controllable light to reconstruct the expected 3D light field.

Similarly to an integrated imaging system, the multi-projector-type light-field display (MPTLFD) system needs to obtain the encoded projector array image (PIA) through the calculation to control the display ray cast by each projector. For the MPTLFD system, common algorithms can be divided into the light-field coding algorithm based on the virtual 3D model and the light-field coding technology based on parallax images (PIs). Among these, light-field coding based on virtual 3D models is widely used, such as the f-Vision system established by Shunsuke in 2013 [12], and the 360° immersive projector array display system proposed by Liu Xu et al. of Zhejiang University in 2014 [14]. Such coding algorithms first need to model objects, obtain all the depth information of the model in advance, and assign values to display pixels by using the light intensity information at the intersection of the display light and the model. The coding principle is relatively simple, while the innovation work of different teams mainly focuses on the hardware structural design. However, the disadvantage is that the modeling process is more complex.

On the other side, the light-field coding algorithm based on PIs is rarely used. In 2010, Balogh et al. [15] directly used the dense sampling camera to capture the parallax image to establish the one-to-one mapping relationship between the display pixel space and the sampling pixel space. This algorithm does not require complex digital modeling. At the same time, the reconstruction effect can retain the true details of the original scene, such as light and shadows, and is more suitable for the light-field display system that reconstructs the real scene. However, under the condition of sparse sampling, owing to the small amount of sampling information, some display rays cannot match their exact corresponding sample, resulting in large errors of the generated projection image array, which affects the reconstruction effect, as shown in the relevant research on integrated imaging [16].

For this reason, this paper proposes a light-field coding algorithm based on target pixel matching. We determine the imaging pixels of the reduction object point required for a display light at the two nearest cameras, and use the pixels to assign a value to the display light, under the condition of sparse sampling; therefore it is still possible to establish a precise mapping relationship between the display pixel and the sampling pixel to generate a high-quality projection image array to improve the light-field reconstruction effect. The effectiveness of the proposed algorithm was verified by experiments.

In Section 2, we briefly introduce the basic structure of the MPTLFD system and the principle of the traditional light-field coding algorithm applied to the system. In Section 3, we present the theory of the algorithm based on target pixel matching. In Section 4, we demonstrate the effectiveness of the algorithm through experiments and the results are also illustrated.

2. Structure of the MPTLFD System and Its Traditional Light-Field Coding Algorithm

2.1. Structure of the Display End and Capture End of the MPTLFD System

A typical structure of a 3D light-field display and capture system based on a projector array is shown in Figure 1. Every projector is uniformly distributed on the circular arc track with the radius R_P according to the interval angle θ_P, and the included angle between the optical axis of a projector or camera, and the positive direction of the x-$axis$ is called the azimuth angle. The azimuth angle range of the projector track is $[\Psi_{min}, \Psi_{max}]$, and every capture camera is evenly arranged on the radius R_C according to the interval angle θ_C, and the azimuth range is $[-\Phi, \Phi]$. R_P and R_C can be different, but they are set to be equal here for ease of calculation. The optical axis of each projector and camera coincides at the center of the track circle. A holographic diffuser with anisotropic light modulation is located between the projector track and the camera track, and the center of the diffuser coincides with the track center. The projected light of each projector gathers at the diffuser, and the 3D light field is reconstructed through the scattering of the diffuser. As the human eye is more sensitive to the change of parallax in the horizontal direction than in the vertical direction when viewing directly against the screen to use the limited data bandwidth to obtain a better stereo vision experience, the system discards the parallax in the vertical direction and only restores the light-field distribution in the horizontal direction. The light field with a smooth horizontal motion parallax is reconstructed by using the projector array arranged in horizontal arcs and the diffuser with a small horizontal scattering angle, which gives the observer a better stereo vision experience and produces a 3D effect with a larger field angle, higher resolution, and more delicate display effect in the horizontal direction.

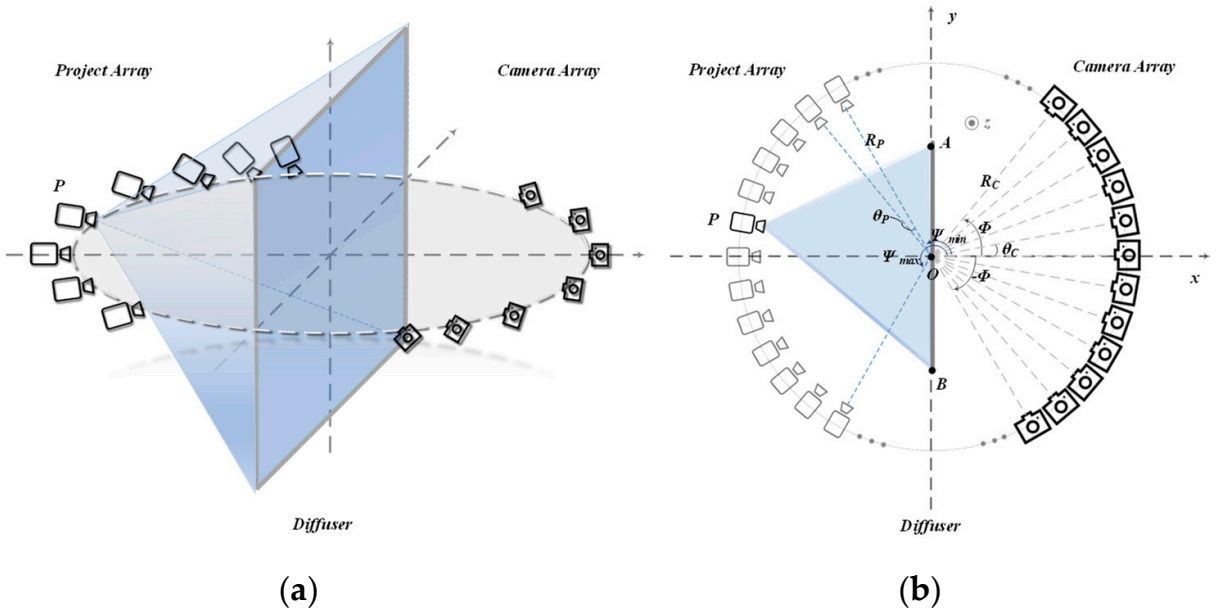


Figure 1. Stereogram (**a**) and top view (**b**) of the capture and display ends.

2.2. The Principle of the Improved SPOC Algorithm Applied to a MPTLFD System

As the projector array in the MPTLFD system casts light similarly to the lens array in the integrated imaging (II) system, the smart pseudoscopic-to-orthoscopic conversion (SPOC) [17] algorithm in the II system can also be used for light-field coding in the MPTLFD system. In the SPOC algorithm, the parameters of the lens array designed for the simulated display end are consistent with those of the capture lens array, so the captured parallax images (PIs) can be used as the elementary image array (EIA) of the simulated display end. To solve the depth inversion problem, it directly used the PI at the simulated display end for optical field coding and synthesized the primitive image array corresponding to the parameters of the real display lens array. When the reference plane in the SPOC algorithm is selected consistently with the plane of the EIA, this step can be equivalent to the direct mapping process from the capture pixel space to the display pixel space.

The SPOC algorithm has become mature after a long-period of development, but it still needs to be appropriately modified when it is applied to the MPTLFD system for coding. The corresponding mapping relationship is modified according to the actual light path, and the algorithm is improved to adapt to the capture and reconstruction parameters of the system. The principle of the improved algorithm applied to the system is shown in Figure 2, which is the top view of the system section at the height of the projector array. The projector array and the capture camera array are arranged in concentric arcs and their configuration are shown in Figure 1a. The camera parameters in the capture space and the projector parameters in the projector array are consistent. Line segment AB in Figure 2 represents the top view of the diffuser.

Set

$J \times K$ as the resolution of the capture camera;

θ_{hov} as the horizontal field of view (FOV);

$M \times N$ as the resolution of the projector array;

For a pixel P_n projected by a projector $P(x_P, y_P)$, use l_{PP_n} represents the display ray between P and P_n, and its geometric expression is:

$$y = \frac{y_P - y_n}{x_P} \cdot x + y_n \tag{1}$$

which is recorded as

$$y = k_d \cdot x + b_d \tag{2}$$

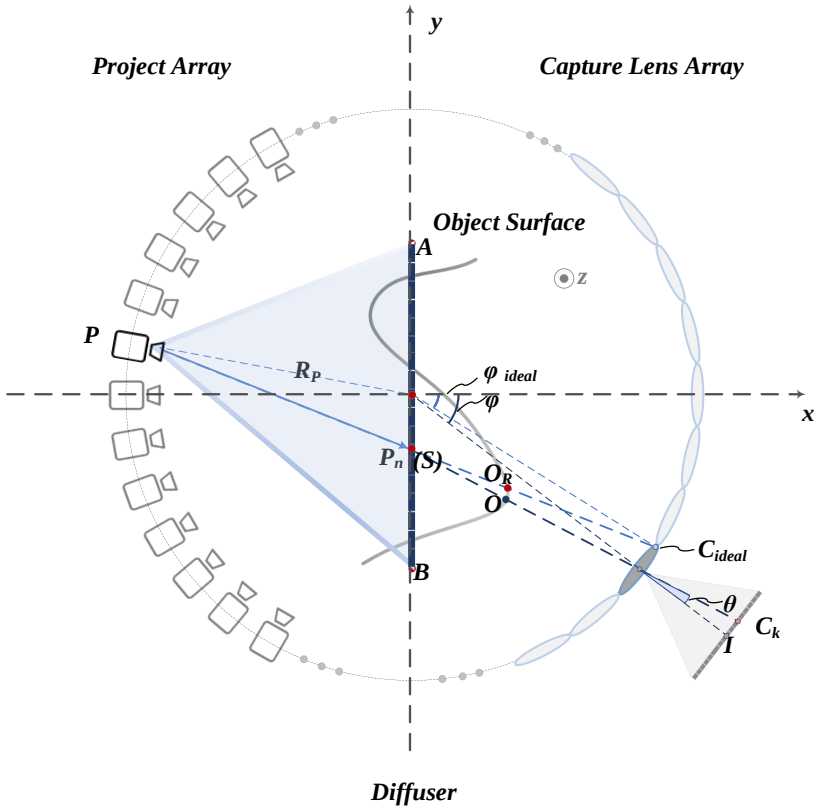


Figure 2. Schematic diagram of the improved SPOC algorithm applied to the MPTLFD system.

Combining the orbit equation of the capture camera $x^2 + y^2 = R_C^2$, we can obtain the coordinate of the ideal camera (C_{ideal}) and then calculate its azimuth φ_{ideal}. However, owing to the discrete capture of the capture camera, there is often no capture camera exactly corresponding to C_{ideal}, in which case the nearest camera is selected; its serial number is C, and its azimuth is φ, where:

$$C = round[(\varphi_{ideal} + \phi)/\theta_C] \tag{3}$$

Taking the plane of the diffuser as the reference plane in the SPOC algorithm, we mark the intersection point of the display ray and the reference plane as S, where S coincides with P_n. The nearest camera (C) is used to capture S to approximate the capture of P_n by C_{ideal}, which is combined with the FOV of the capture camera (θ_{hov}) to determine pixel C_k by Equation (4) corresponding to the capture ray between C and S (l_{CS}) on the imaging plane (I):

$$C_k = round\left\{\frac{K}{2} + \frac{\tan(\theta)}{\tan\left(\frac{\theta_{hov}}{2}\right)\Big/\left(\frac{K}{2}\right)}\right\}, \theta = arctan(k_d) - \varphi \tag{4}$$

where θ represents the angle between the optical axis of C and the capture ray l_{CS}.

In fact, the system only restores the horizontal motion parallax of the object, while in the vertical direction, the divergence angle of the diffuser is large enough to ensure that 3D objects can be observed only in a wide range of vertical field angles, which also causes the mapping relationship of the coding algorithm to remain unchanged at each vertical height. Therefore, the mapping relationship between the capture column pixels and the display column pixels can be established according to Equation (5) to calculate each column of pixels projected by each projector in turn. Finally, the correct projector image array (PIA) is obtained.

$$P_n^m = C_k^m, m \in [0, M-1] \tag{5}$$

where P_n^m represents the value of the pixel in the mth column and P_nth row of the Pth projector image and C_k^m represents the value of the pixel in the mth column and C_kth row of the Cth capture image.

In the experiment, we found that when the camera spacing was too large, the quality of the PIA using the SPOC coding algorithm decreased, resulting in a poor light-field reconstruction effect. This is because, as shown in Figure 2, the display ray l_{PP_n} is expected to restore the information of object point O_R, while in the SPOC algorithm, the nearest camera's corresponding capture ray is used to replace the ideal camera to assign a value to the display ray, and the capture ray actually captures the light-field information at object point O, thus introducing an error. At the same time, the error increases with the increase of the capture camera's arrangement interval.

3. Sparse Light-Field Coding Algorithm Based on Target Pixel Matching

To solve the error in the SPOC algorithm under sparse capture conditions, this paper proposes an improved light-field coding algorithm based on target pixel matching. Considering that the difference of the same object point's light intensity captured by adjacent cameras is small, by searching for the pair of most similar pixel columns in the two cameras nearest to the display ray, the algorithm can obtain the correct imaging pixel columns of the actual object, which are then used to assign values to the display ray. This algorithm can establish a more accurate mapping relationship between the capture pixel and the display pixel under the condition of low capture density, synthesize a high-quality PIA, and significantly improve the light-field reconstruction effect.

As shown in Figure 3**,** the optical center of the two nearest capture cameras C_A, C_B can be obtained by Equation (6), and obtains the camera azimuth φ_A, φ_B, where:

$$C_A = \left\lfloor \frac{\varphi_{idedal} + \phi}{\theta_C} \right\rfloor, \; C_B = C_A + 1 \tag{6}$$

the object point O_R expected to be restored for the display ray l_{PP_n} is located on the line segment between the point P and the point C_{ideal}. Record the imaging pixels of O_R captured by C_A and C_B as I_{O_A}, I_{O_B}, and they can be, respectively, obtained by calculating the intersection of the imaging plane and the capture ray from O_R to the optical center of the camera. In other words, O_R is located at the intersection of the display ray and the capture ray from the optical center of the camera to the imaging pixel of O_R. Therefore, this algorithm first determines the interval where I_{O_A} exists in C_A, for any pixel within the interval, a capture ray can be determined from the pixel to the optical center of the camera. As mentioned above, O_R is located at the intersection of the display light and a certain capture ray, according to the intersection, the corresponding imaging pixel in C_B can be determined, and multiple pixel pairs can be established by traversing all pixels within the interval. Generally, an object can be regarded as a Lambert body, and the light emitted by the same object is isotropic at all angles. In fact, the distance between the object and the capture camera is far greater than the optical center interval of the camera, when the adjacent cameras capture the same object point, the included angle of the capture rays is small, and the difference of the capture rays intensity information is small, the values of the imaging pixels camera are similar, so, I_{O_A} and I_{O_B} are the group of pixel pairs with the highest similarity. The schematic diagram of the algorithm is shown in Figures 4 and 5.

To reduce the computational redundancy, the algorithm first establishes a virtual activity space, which is a cylinder space with a sufficiently large radius, and the space can contain all of the light-field information of the real scene. As shown in Figure 4, this figure is the top view of the system structure at the height of the projector array, the blue area with a radius R_A is the virtual activity space. l_{PP_n} intersects the boundary of the active space at two points, which are recorded as the near field point, $Near(x_N, y_N)$, and far field point, $Far(x_F, y_F)$ according to the distance between these two points and the camera's orbit.

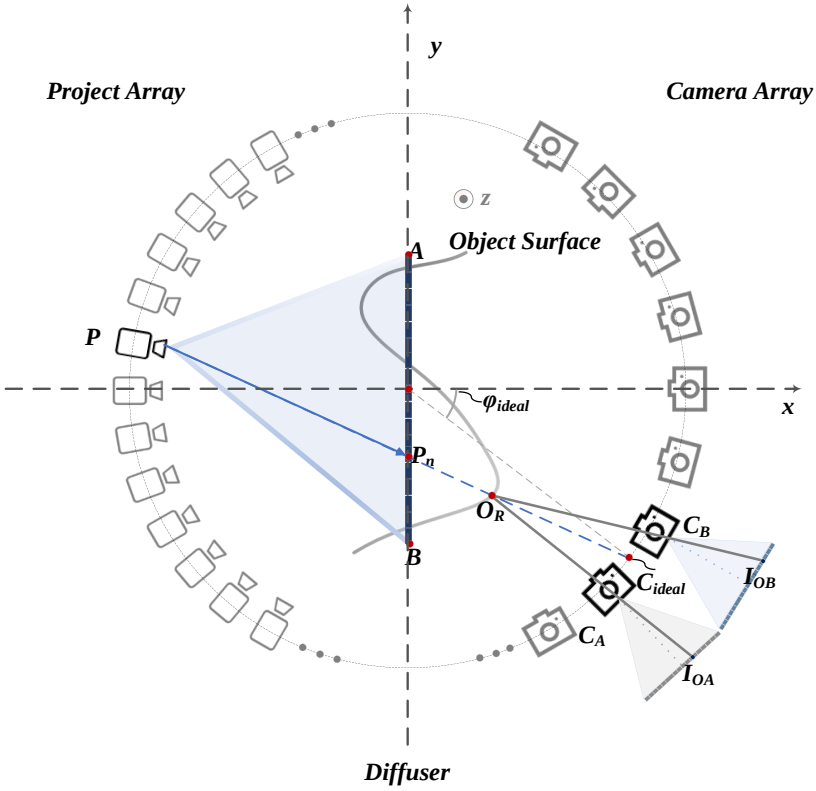


Figure 3. The two nearest cameras and the imaging pixels of O_R.

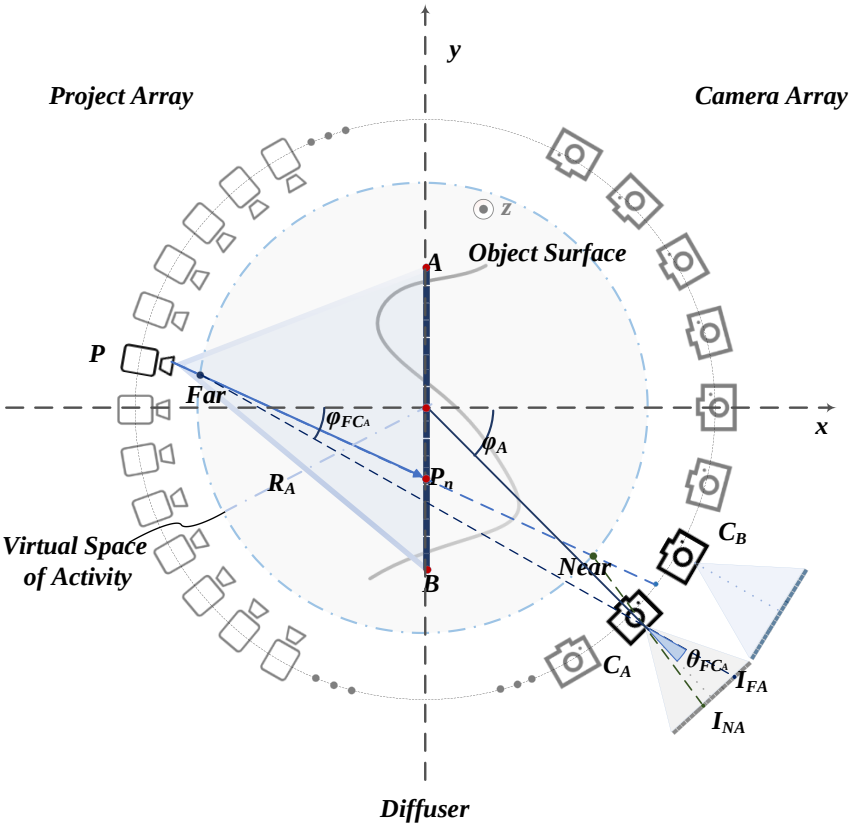
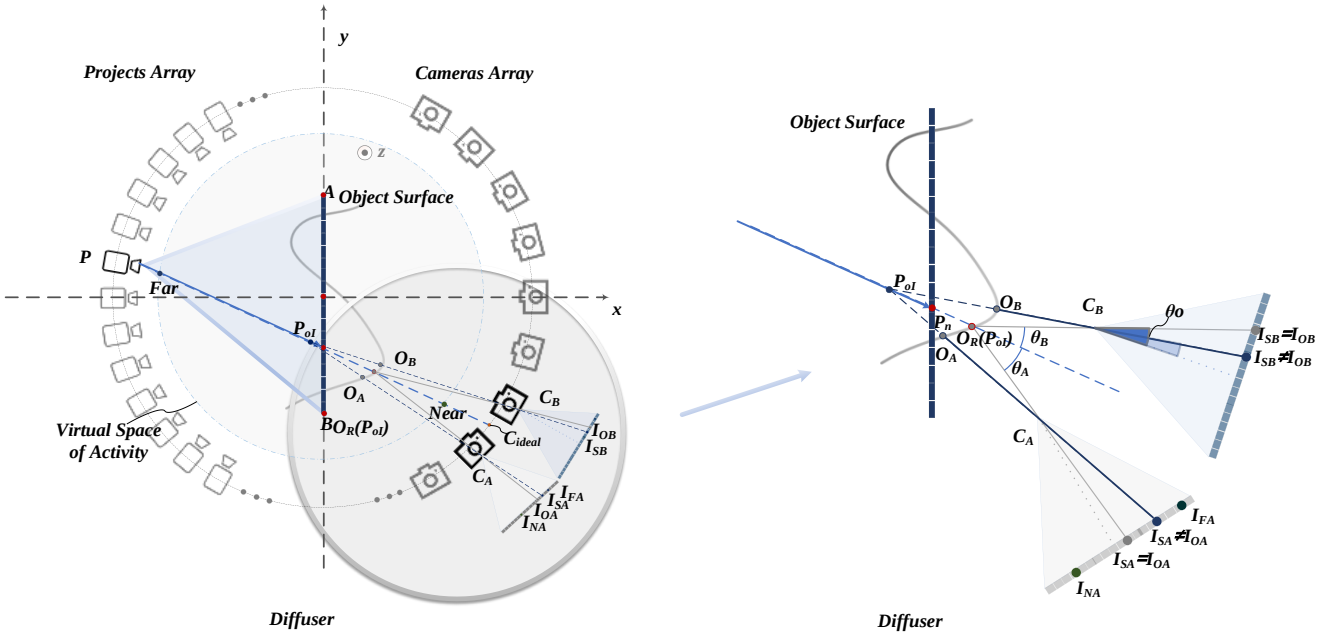


Figure 4. Virtual activity space and interval of the imaging pixels.



(**a**) (**b**)

Figure 5. Schematic diagram of light-field coding algorithm based on target pixel matching (**a**) and local enlargement (**b**).

According to the relationship between the far and near field points and the optical center C_A, and in combination with the FOV of the camera, the algorithm calculates the imaging pixels I_{F_A}, I_{N_A}, where:

$$I_{F_A} = round\left\{\frac{K}{2} + \frac{\tan(\theta_{FC_A})}{\tan\left(\frac{\theta_{hov}}{2}\right)\Big/\left(\frac{K}{2}\right)}\right\}, \theta_{FC_A} = \varphi_{FC_A} - \varphi_A \tag{7}$$

$\theta_{FC_A}, \varphi_{FC_A}$ represent the angle between the capture ray l_{FC_A} and C_A and the *x-axis*, respectively; then I_{N_A} can be obtained in the same way.

The real object point (O_R) to be restored for l_{PP_n} is located in the area of the line segment (l_{FA}) between *Near* and *Far*, so I_{O_A} meets the following relationship:

$$0 \le I_{N_A} \le I_{O_A} \le I_{F_A} \le K - 1 \tag{8}$$

Equation (8) determines the interval where the imaging pixel of O_R exists and then determines the actual imaging pixel of O_R. Assume that the imaging pixel of O_R in camera C_A is I_{S_A}, and $I_{S_A} \in \left[I_{N_A}, I_{F_A}\right]$. Then, the angle between the capture ray and the *x-axis* can be deduced by Equation (7), and then the mathematical expression can be obtained, which is recorded as:

$$y = k_C \cdot x + b_C \tag{9}$$

The coordinate of intersection P_{oI} of the capture ray from I_{S_A} to $C_A(l_{I_{S_A}C_A})$ and l_{PP_n} is as follows:

$$P_{o_I}\left(x_{P_{o_I}}, y_{P_{o_I}}\right), \begin{cases} x_{P_{o_I}} = (k_C - k_d)/(b_d - b_C) \\ y_{P_{o_I}} = k_C \times x_{P_{o_I}} + b_C \end{cases} \tag{10}$$

Then, O_R is at P_{oI}. For the adjacent camera C_B, the imaging pixel of O_R is the intersection point of the imaging plane and the capture ray from P_{oI} to $C_B(l_{P_{o_I}C_B})$, and then the imaging pixel is recorded as I_{S_B}:

$$I_{S_B} = round\left\{\frac{K}{2} + \frac{\tan(\theta_O)}{\left[\tan\left(\frac{\theta_{hov}}{2}\right)\Big/\left(\frac{K}{2}\right)\right]}\right\}, \theta_O = \varphi_O - \varphi_B \tag{11}$$

where θ_O, φ_O represent the angle between the capture ray $l_{P_{o_I}C_B}$ and the optical axis of C_B and the $x-axis$, respectively. According to the geometric relationship, multiple pixel pairs are obtained by traversing all pixels within the interval $\left[I_{N_A}, I_{F_A}\right]$.

As shown in Figure 5b, when $I_{S_A} \neq I_{O_A}$, capture rays $l_{I_{S_A}C_A}$, $l_{P_{o_I}C_B}$ and the real object surface intersect at O_A, O_B, respectively, it actually represents that I_{S_A} is the imaging pixel of O_A by camera C_A, while I_{S_B} is the imaging pixel of O_B by camera C_B. O_A is completely unrelated to O_B, so the imaging pixels are different. When the assumed imaging pixel I_{S_A} is the real imaging pixel I_{O_A}, I_{O_A} and I_{O_B} represent the imaging pixels of the same real object point O_R in the two nearest cameras, and the pixel values are similar at this time. To simplify the calculation, we use the variance value to represent the similarity of the pixels, and the smaller the variance, the more similar the pixels. According to the above analysis, the imaging pixels of the real object point in the two nearest cameras can be given by traversing the pixel pairs with geometric correspondence and obtaining the pixel pair with the smallest variance. Similarly, as mentioned above, the mapping relationship of the light-field coding algorithm does not change at each vertical height, so the object point information O_R actually represents a continuous list of object point information, of which the light field is captured dispersed into J pixels, where J represents the camera's vertical resolution, so the imaging pixel columns of the real object point O_R at the two nearest cameras can be calculated by Equation (12) when $s^2 = \sum_{j=0}^{J-1} \frac{(I^j_{S_A} - (I^j_{S_A} + I^j_{S_B})/2)^2 + (I^j_{S_B} - (I^j_{S_A} + I^j_{S_B})/2)^2}{2}$ is minimum.

$$I_{O_A} = I_{S_A}, I_{O_B} = I_{S_B} \tag{12}$$

Here, $I^j_{S_A}$ and $I^j_{S_B}$, respectively, represent the value of the jth row of column pixel I_{S_A} in the C_Ath captured image and the value of the jth row of column pixel I_{S_B} in the C_Bth captured image.

The light emitted from the same object point to different angles has a correlation. The smaller the angle between the lights, the greater the correlation. In the process of pixel synthesis, the smaller the angle between the capture ray and the display ray, the greater the impact of the capture pixel on the display pixel. Therefore, the capture pixel is weighted and fused into the pixel value of the PIA by interpolation according to the mapping relationship in Equation (13). Calculate each column of pixels projected by each projector in turn, and finally obtain the correct PIA.

$$P^m_n = \Delta \cdot I^m_{O_A} + (1-\Delta) \cdot I^m_{O_B}, \Delta = \theta_B/(\theta_A + \theta_B) \tag{13}$$

where Δ is the weight; θ_A, θ_B are the angles between the two capture rays and the display ray l_{PP_n}; P^m_n, $I^m_{O_A}$, $I^m_{O_B}$ represent the pixel value of the mth row pixel in the display pixel column P_n and capture pixel column I_{O_A}, I_{O_B} , respectively.

This approach can be summarized as follows:

(1) Calculate the two capture cameras that are closest to a display ray cast by the projector P;
(2) Determine the interval of the imaging pixel of the object that the display ray needs to restore;
(3) Determine the coordinate of the intersection of the display ray and the line between a pixel in the interval and the optical center of the camera;
(4) Calculate the image pixel of the intersection coordinate taken by another nearest camera;
(5) Traverse all pixels in the interval to obtain multiple pixel pairs with the above relation;
(6) Find the pixel pair that minimizes s^2 to determine the imaging pixels of the object;
(7) Calculate the display pixels values by Equation (13).

4. Experimental Verification and Analysis

The display system used in the experiment consisted of 108 projectors with a resolution of $1280 * 720$ and a holographic diffuser. The projectors were uniformly distributed on the circular track with a radius of 2.14 m and an azimuth of [153°, 207°] according to an interval angle of 0.5°. The center of the diffuser coincided with the center of the projector track and was perpendicular to the optical axis of the central projector (with an azimuth angle of 180°). The specific display system structure is shown in Figure 6.

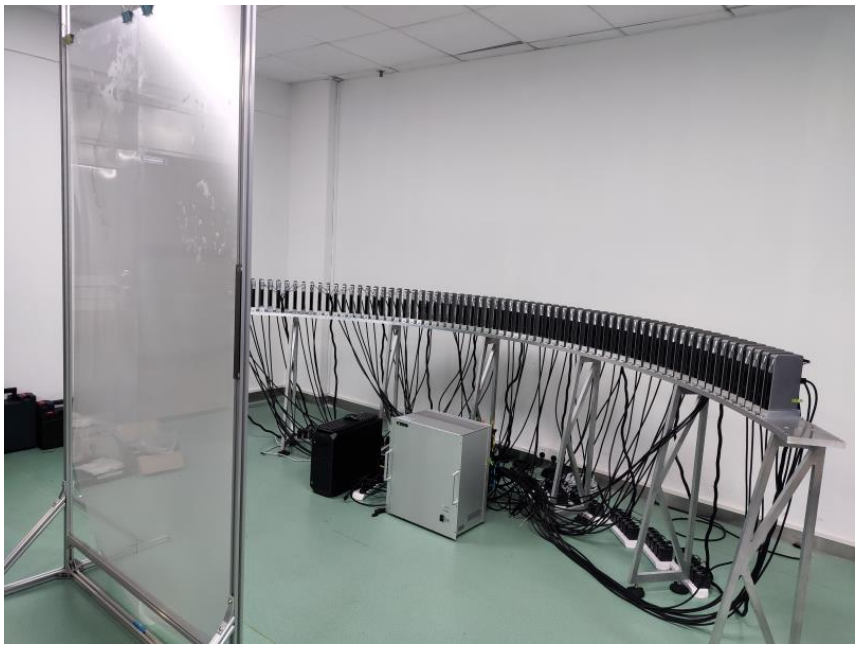
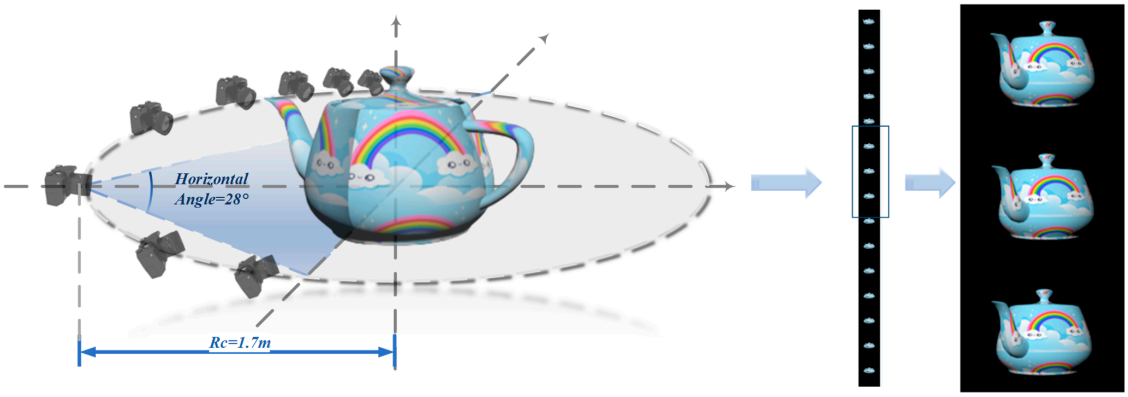


Figure 6. Picture of the real display system structure.

4.1. Structure of the Virtual Scene Capture End

To verify the effectiveness of the algorithm, first, we selected the 3DMax software to capture the 3D light-field information. The structure and scene settings of this capture system are shown in Figure 7a. The FOV of the virtual camera was28°, and the capture image resolution was$1280 * 720$. The cameras were evenly distributed on the circular arc track, with a radius of 2.14 m and an azimuth of [−48.6°, 48.6°], according to a 1.8° spacing angle. There were 55 cameras in total, and the optical axes of each camera converged at the center of the track. The captured object was a teapot with a cartoon sticker. The geometric center of the teapot coincided with the center of the camera track.

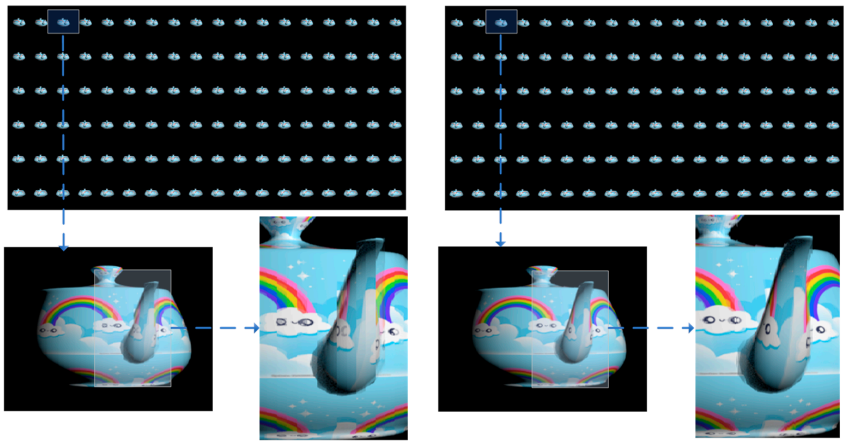


(**a**) (**b**)

Figure 7. Scene setting of the virtual acquisition terminal (**a**) and PIs obtained by capture (**b**).

4.2. Experimental Results and Analysis

Following capturing the sparse PIs by the above method, we compare the PIA generated by the proposed algorithm with the PIA generated by the SPOC algorithm, and the results are shown in Figure 8. On the whole, the two PIAs are similar in terms of image structure, teapot contour, cartoon sticker patterns, and the distribution patterns. For the shadow areas caused by uneven lighting, the two images have similar rendering effects. However, the details are different. The quality of the PIA is better when using the target pixel matching algorithm, and the layout of the whole image in Figure 8b is fine and smooth, without the severe stripe mosaic phenomenon shown in Figure 8a. Compared with the original PIs, such as the eyeball of a cartoon sticker, the lines shown in Figure 8a are disordered, while those in Figure 8b are even, which is close to the original scene. For the part with complex depth information in the original scene, the spout in Figure 8a has a severe ghosting phenomenon, and it is difficult to distinguish the depth change relationship between the spout and the body. The spout in Figure 8b is smooth and round, and there is a distinct boundary between the spout and the body. This is because in the SPOC algorithm, it is difficult to select a single reference plane to reflect the complex depth information of the continuous changes of objects. In this algorithm, the reference plane is no longer introduced manually, and the similarity comparison of target pixels is used to obtain the imaging pixels of real objects and assign values to the PIA. Under the sparse capture condition, the imaging information of objects at different depths can be restored more accurately.



(**a**) (**b**)

Figure 8. Overall and local enlargement of the projector image array. The PIA generated by the SPOC algorithm (**a**) and target pixel matching algorithm (**b**).

The peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), which are commonly used in image quality evaluation, are introduced to test the performance of the proposed algorithm. The PIA generated by capturing PIs under dense capture conditions was used as the base image, and the PIAs generated by the two algorithms under the condition of sparse capture as experimental images. PSNR and SSIM were calculated in the following two cases:

case 1: compare the first single projection image of the base image with the first single projection image of the experimental images;

case 2: compare all the projection images of the base image with all of the projection images of the experimental images.

The results are shown in the Tables 1 and 2, which also compare the encoding time of the two algorithms (the time of calculating a single projection image from the PIs through the optical field-coding algorithm). For whether a single projection image or all PIAs, the PSNR and SSIM values of the images generated by this algorithm are higher, a higher PSNR means a smaller distortion rate compared with the base image, a higher SSIM indicates that the experimental image has a higher similarity with the base image in contrast, brightness, structure and other aspects. In short, PIA generated by this algorithm are more similar to the original scene, proving the effectiveness of the proposed light-field encoding algorithm.

Table 1. Comparison of PSNR and SSIM of the two algorithms in case 1.

Algorithm	PSNR(dB)	SSIM	Encoding Time(S)
SPOC	38.68	0.65	15.69
Target pixel matching	39.93	0.82	120.24

Table 2. Comparison of PSNR and SSIM of the two algorithms in case 2.

Algorithm	PSNR(dB)	SSIM	Encoding Time(S)
SPOC	39.13	0.72	15.69
Target pixel matching	40.76	0.88	120.24

However, the encoding time of the proposed algorithm is longer, the algorithm speed is mainly limited by the number and resolution of projectors and the selection of the radius of the virtual activity space. When the projector interval is increased or its horizontal resolution is reduced, the algorithm encoding time is also reduced accordingly, but in the display effect, the motion parallax in the horizontal direction is not smooth enough. When the radius of the virtual activity space is reduced, because the distance between the far and near field points is reduced, the determined imaging interval in a camera becomes smaller, and the time cost is also reduced when searching for the most similar pixel pairs by traversing all possible imaging pixels in the imaging interval. However, when the virtual activity space is too small to contain all of the light-field information of the object, all of the information of the original 3D scene cannot be reconstructed. In the experiment, affected by the system structure, the PIA was synthesized pixel column by pixel column. The encoding process of each column of display pixels is independent and irrelevant, so the algorithm speed can be improved through a parallel operation.

We loaded the PIAs encoded by the SPOC algorithm and the algorithm proposed in this paper into the MPTLFD system to observe the actual display effect, and the comparison of the display results is shown in Figure 9. It can be seen that the SPOC algorithm has a poor reconstruction effect, severe color aliasing, fuzzy cartoon stickers, and an irregular arrangement of stripe mosaics. On the contrary, in the reconstructed light-field effect based on the target pixel matching algorithm, first, as the perspective changes, the reconstructed object shows smooth and delicate different sides and has complete and continuous horizontal motion parallax and a good stereo sense. In addition, in the SPOC algorithm, in the left 30° view, it is obvious that the right side of the reconstructed object has a phantom

beyond the contour of the teapot; in the middle view, the spout and the body overlap, making them almost indistinguishable, and the phenomenon of ghosting is serious. In contrast, in the algorithm proposed in this paper, the contour of the teapot is regular, the boundary between the teapot spout and the teapot body is clear, and the depth information is different, which is more similar to the original scene. This is because for each display ray observed from a certain perspective, it restores different object point information. Owing to the lack of capture information caused by sparse capture, the SPOC algorithm directly uses the nearest camera instead of the ideal camera to assign values to the display ray. As shown in Figure 2, the display ray l_{PP_n} that needs to restore the light-field information of O restores the information of O_I by error, the imprecise mapping relationship causes numerous display rays to carry the wrong and the same object information. From the perspective of the observer, they can observe the ghosting phenomenon. At the same time, for the teapot spout, a single reference plane introduced in the SPOC algorithm cannot accurately represent its complex depth information, resulting in no obvious change in the depth of the spout and the body in the final display effect. In this algorithm, according to the target pixel, which is the actual imaging pixel of the object point information that each display ray needs to restore, the display ray is assigned a value to ensure that each it can restore accurate object point information, through this way, this algorithm can effectively solve the ghosting phenomenon, accurately restore the information at different depths, and improve the 3D effect.

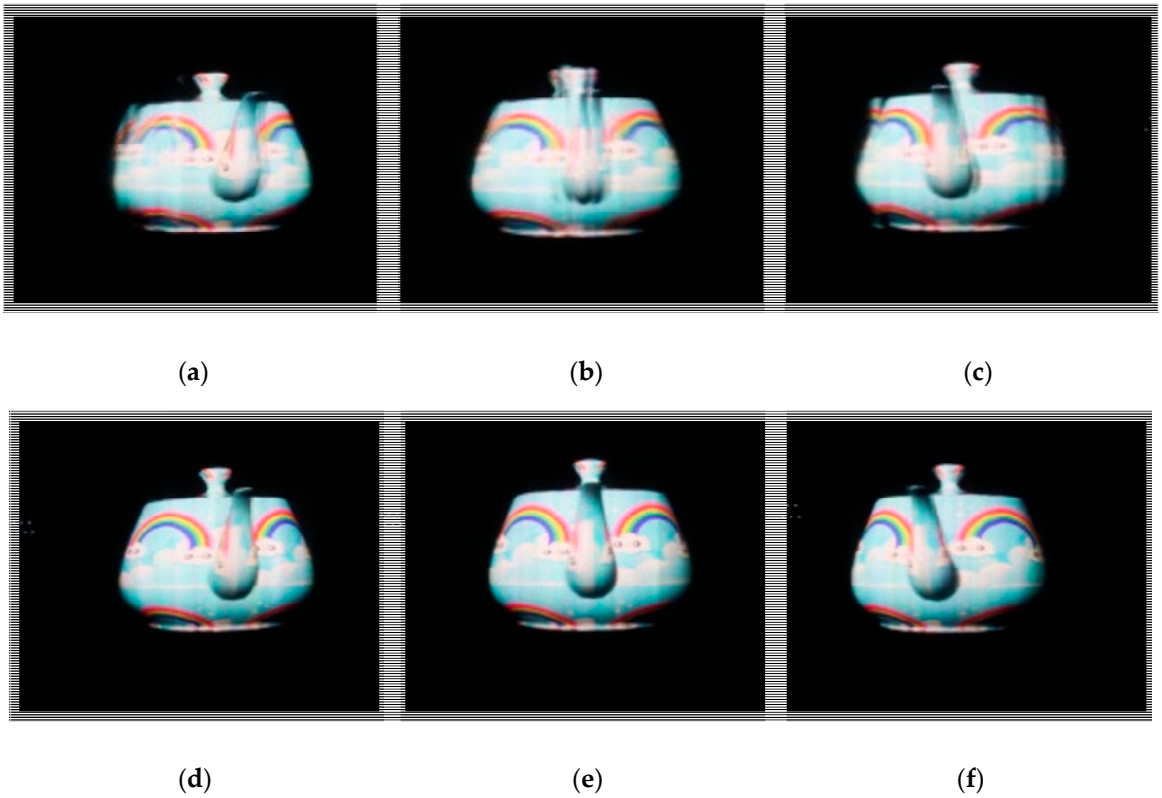


Figure 9. Light-field reconstruction effect under different viewing angles. SPOC algorithm: (**a**) left 30° view, (**b**) middle view, (**c**) right 30° view. Target pixel matching algorithm: (**d**) left 30° view, (**e**) middle view, (**f**) right 30° view.

Through the above experiments, we confirmed that the proposed algorithm could accurately synthesize the PIA under sparse capture conditions, effectively improve the effectiveness of the light-field reconstruction for complex scenes with depth information, and is more suitable for high-quality light-field display systems.

5. Conclusions

In this paper, a new light-field coding algorithm for synthesizing a PIA was proposed. Under the condition of sparse capture, the PIs obtained by the capture is directly used. First, the auxiliary virtual activity space is set to determine the interval where the imaging pixel of the object exists, and the imaging pixel of the object under different depth information can be accurately determined by finding the most similar pixel pair in the two nearest cameras. The imaging pixel is used to assign a value to the display pixel to establish a more accurate mapping relationship between the capture pixel space and the display pixel space, thereby improving the image synthesis quality. Compared with the model-based light-field coding algorithm, this algorithm does not need a complex modeling process and is more concise and efficient. Compared with the SPOC algorithm, this algorithm does not need to introduce a reference plane, which avoids the problem of introducing errors by using a single reference plane to reflect the continuous changes in the depth information of the actual 3D scene. Through experimental comparison, we found that the PIA synthesized by the algorithm proposed in this paper was clearer and more accurate, and the reconstructed light field was more similar to the original scene after being loaded into the display system. However, this algorithm also has some problems, such as a long encoding time. At the same time, in the face of the problem of light-field reconstruction of the real scene, in addition to the problem of sparse camera arrangement, the camera attitude should also be considered. Subsequent work will focus on the acceleration of the light-field encoding algorithm and the coding of free cameras in a real scene.

Author Contributions: Conceptualization, X.J. and X.S.; methodology, Q.M. and X.J.; software, Q.M. and H.Y.; validation, Q.M. and X.J.; data curation, Q.M. and H.Y.; writing—original draft preparation, Q.M. and H.Y.; writing—review and editing, Q.M. and H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Beijing Municipal Science & Technology Commission, Administrative Commission of Zhongguancun Science Park No. Z221100006722023 and National Natural Science Foundation of China (62085016).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Son, J.-Y.; Lee, H.; Lee, B.-R.; Lee, K.-H. Holographic and Light-Field Imaging as Future 3-D Displays. *Proc. IEEE* **2018**, *105*, 789–804. [CrossRef]
2. Wang, Z.; Lv, G.; Feng, Q.B.; Wang, A.T.; Ming, H. Resolution priority holographic stereogram based on integral imaging with enhanced depth range. *Opt. Express* **2019**, *28*, 2689–2802. [CrossRef] [PubMed]
3. Zhang, X.; Lv, G.; Wang, Z.; Hu, Z.; Ding, S.; Feng, Q. Resolution-enhanced holographic stereogram based on integral imaging using an intermediate-view synthesis technique. *Opt. Commun.* **2020**, *458*, 124656. [CrossRef]
4. Mao, Y.; Wang, W.F.; Jiang, X.Y.; Zhang, T.; Yu, H.Y.; Li, P.; Liu, X.L.; Le, S. Elemental image array generation algorithm with accurate depth information for integral imaging. *Appl. Opt.* **2021**, *60*, 9885–9886. [CrossRef] [PubMed]
5. Sakamoto, K.; Morii, T. *Development of the 3D Adapter Using an Optical Grating Film for Stereoscopic Viewing*; SPIE Optics East: Boston, MA, USA, 2006.
6. Luo, J.-Y.; Wang, Q.-H.; Zhao, W.-X.; Tao, Y.-H.; Li, D.-H. Autostereoscopic Three Dimensional Display Based on Two Parallax Barriers. In Proceedings of the 2010 Symposium on Photonics and Optoelectronics, Chengdu, China, 19–21 June 2010.
7. Kao, Y.Y.; Huang, Y.P.; Yang, K.X.; Chao, P.C.P.; Tsai, C.C.; Mo, C.N. 11.1: An Auto-Stereoscopic 3D Display Using Tunable Liquid Crystal Lens Array That Mimics Effects of GRIN Lenticular Lens Array. *Sid Symp. Dig. Tech. Pap.* **2009**, *40*, 111–114. [CrossRef]
8. Blanche, P.A.; Bablumian, A.; Voorakaranam, R.; Christenson, C.; Lin, W.; Gu, T.; Flores, D.; Wang, P.; Hsieh, W.Y.; Kathaperumal, M.; et al. Holographic three-dimensional telepresence using large-area photorefractive polymer. *Nature* **2010**, *468*, 80–83. [CrossRef] [PubMed]
9. Matsushima, K.; Sonobe, N. Full-color digitized holography for large-scale holographic 3D imaging of physical and nonphysical objects. *Appl. Opt.* **2018**, *58*, A150–A156. [CrossRef] [PubMed]

10. Zhang, Y.; Fan, H.; Wang, F.; Gu, X.; Qian, X.; Poon, T. Polygon-based computer-generated holography: A review of fundamentals and recent progress [Invited]. *Appl. Opt.* **2022**, *61*, B363–B384. [CrossRef] [PubMed]
11. Takaki, Y.; Takenaka, H.; Morimoto, Y.; Konuma, O.; Hirabayashi, K. Multi-view display module employing MEMS projector array. *Opt. Express* **2012**, *20*, 28258. [CrossRef] [PubMed]
12. Yoshida, S. Real-time rendering of multi-perspective images for a glasses-free tabletop 3d display. In Proceedings of the 2013 3DTV Vision Beyond Depth (3DTV-CON), Aberdeen, UK, 7–8 October 2013.
13. Jones, A.; Nagano, K.; Liu, J.; Busch, J.; Yu, X.; Bolas, M.T.; Debevec, P.E. Interpolating Vertical Parallax for an Autostereoscopic 3D Projector Array. *J. Electron. Imaging* **2014**, *23*, 011005.
14. Zhong, Q.; Chen, B.S.; Li, H.F.; Liu, X.; Xia, J.; Wang, B.P.; Xu, H.S. Multi-projector-type immersive light field display. *Chin. Opt. Lett.* **2014**, *12*, 060009. [CrossRef]
15. Balogh, T.; Kovács, P.T. Real-time 3D light field transmission. In Proceedings of the Real-Time Image and Video Processing 2010, San Jose, CA, USA, 16 April 2010.
16. Huang, Y.Q.; Yan, Z.; Jiang, X.Y.; Jing, T.; Chen, S.; Lin, M.; Zhang, J.G.; Yan, X.P. Performance Enhanced Elemental Array Generation for Integral Image Display Using Pixel Fusion. *Front. Phys.* **2021**, *9*, 639117. [CrossRef]
17. Navarro, H.; Martínez-Cuenca, R.; Saavedra, G.; Martínez-Corral, M.; Javidi, B. 3D integral imaging display by smart pseudoscopic-to-orthoscopic conversion (SPOC). *Opt. Express* **2010**, *18*, 25573–25583. [CrossRef] [PubMed]