LASER & PHOTONICS REVIEWS
www.lpr-journal.org

Super-Large Field-of-View, High-Accurate and Real-Time 3D Scene Reconstruction Based on Metasurface-Enabled Structured Light

Zhengren Zhang, Anjun Qu, Mengran Yang, Zile Li,* Tian Huang, Shuqi Chen, Hua Cheng, Shaohua Yu, and Guoxing Zheng**

Three-dimensional (3D) reconstruction based on structured light (SL) has wide applications in fields such as artificial intelligence, surveying and mapping, and precision manufacturing. An ideal 3D reconstruction system features both wide field-of-views (FOV), high-accuracy and high-speed reconstruction and ultracompact volume. However, these criteria have been achieved only partially with conventional SL techniques like optical projector and differentiate optical element, which will substantially limit their practical applications. In this study, an ultracompact SL and its 3D reconstruction system is designed and experimentally demonstrated, which exhibits super-large FOV angles up to 210° × 210°, only 3.8‰ reconstruction errors, real-time processing speeds up to 30 fps and SL dimensions as small as 0.3 mm × 0.3 mm. Specifically, the SL is realized with a resonant geometric metasurface, which can produce tens of thousands of random dots and project them into the whole 4π free space. The metasurface enabled SL is collaborated with a binocular stereoscopic camera and advanced 3D reconstruction algorism to form a practical 3D reconstruction system, which can provide robust hardware and software support for high-accuracy and real-time 3D scene reconstruction.

1. Introduction

Three-dimensional (3D) scene reconstruction technology, with its high authenticity, has emerged as a critical foundation for numerous applications, including human-robot interaction, virtual reality (VR) and automobile manufacturing.[1] In the existing technological landscape, 3D scene reconstruction predominantly relies on three principal methods: LIDAR reconstruction,[2–4] visual reconstruction,[5,6] and structured light (SL)-based reconstruction.[7–9] The LIDAR reconstruction method employs a LIDAR system to scan the environment and measure the time of flight interval of laser pulses from emission to reception, calculating 3D data like distance and angle of target.[10] However, this method is limited by the rotational scanning mechanisms and the requirement for dense pulse emissions,[11] resulting in a complex and bulky system structure. These factors impede the system's progress in miniaturization and precision, presenting challenges in developing compact, integrated devices. The visual reconstruction method captures multi-view images of the target scene through the camera, extracts depth values using the geometric relationship between the image and the real space, and obtains the 3D information of the entire scene via feature matching and point cloud stitching.[12] Given that the image taken by the camera in a single-shot is inherently 2D, achieving a broader field-of-view (FOV) is more straightforward in visual reconstruction than with the 1D pulse of LIDAR. The larger the FOV, the more 3D information the image can contain, increasing the information capacity in reconstruction results. However, cameras are sensitive to lighting conditions. Excessive or insufficient illumination can induce computational errors in the system or even result in failure to obtain depth information. Furthermore, for scenes lacking visual features, purely visual approach based on texture information in images struggle to achieve reliable reconstruction results. In recent years, SL-based 3D reconstruction technology has made significant progress due to the development of diffractive optics and on-chip image processing. This technology can effectively reduce the interference of ambient light and actively endow the

Z. Zhang, A. Qu, M. Yang, S. Yu
School of Materials Science and Engineering
Chongqing Jiaotong University
Chongqing 400074, China
E-mail: zzr@cqjtu.edu.cn

Z. Li, G. Zheng
Peng Cheng Laboratory
Shenzhen 518055, China
E-mail: lizile@whu.edu.cn; gxzheng@whu.edu.cn

Z. Li, G. Zheng
Wuhan Institute of Quantum Technology
Wuhan 430206, China

Z. Li, T. Huang, G. Zheng
Electronic Information School, and School of Microelectronics
Wuhan University
Wuhan 430072, China

S. Chen, H. Cheng
The Key Laboratory of Weak Light Nonlinear Photonics, Ministry of Education, School of Physics and TEDA Institute of Applied Physics
Nankai University
Tianjin 300071, China

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/lpor.202401120

DOI: 10.1002/lpor.202401120

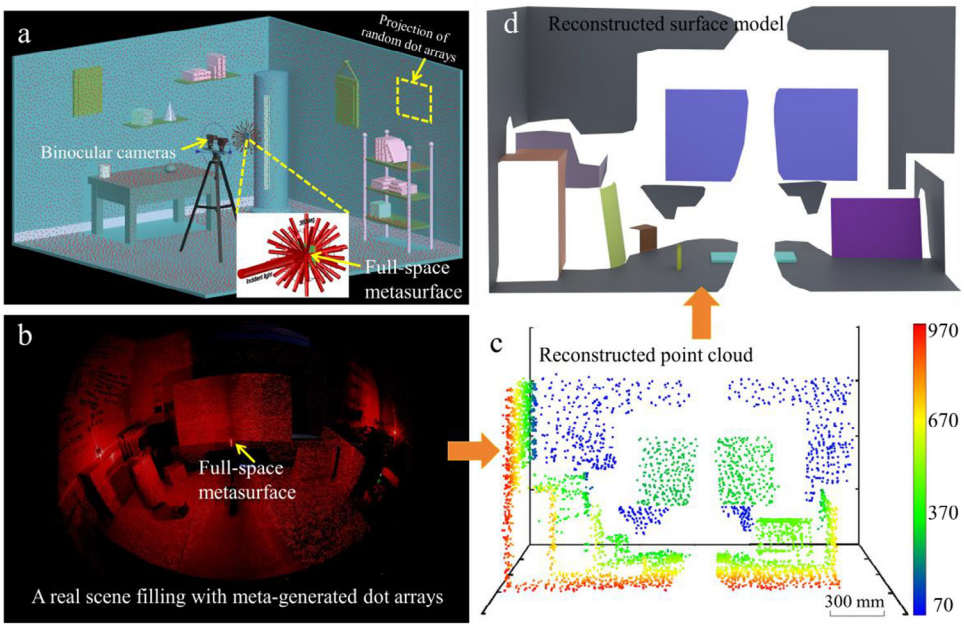

Figure 1. 3D scene reconstruction using metasurface-projected random dot arrays. a) The full-space metasurface projects random dot arrays across the entire 4π free space, marking the objects in the space uniquely. The reflected images are captured through scanning with a self-construct binocular camera, utilizing the principle of triangular similarity and a real-time mapping algorithm to obtain the 3D point cloud of objects within the scanned area, facilitating the reconstruction of the scene model. b) A real scene filling with meta-generated dot arrays and its reconstruction results shown in with point cloud c) and 3D surface model d), different colors in d represent different objects. Interestingly, because of its ultracompact volume, the metasurface-based SL projector can be put into the center of the scene with very limited occlusion for 3D reconstruction.

target with texture features by integrating dot arrays SL into visual reconstruction.[13] It enables one-shot modeling of surfaces without mechanical scanning, thus compensating for the limitations of the LIDAR method and the purely visual method, while greatly enhancing the efficiency and accuracy of 3D reconstruction. Nonetheless, the phase modulation of conventional optical elements based on Snell's law remains limited in miniaturization due to its large thicknesses, and the coverage of the projected dot arrays SL (FOV is typically lower than 65°) at the current technical level, presenting difficulties in meeting the measurement requirement for broader spaces and integration in 3D reconstruction systems.[14] Therefore, it is urgently necessary to implement new principles and methods to solve the current predicaments.

Metasurfaces are ultra-thin artificial surfaces composed of subwavelength cell structures arranged periodically or non-periodically, which can exhibit parametric properties that exceed natural materials, and possess a powerful ability to modulate the phase,[15–17] polarization,[18–20] propagation modes,[21–23] and other characteristics of optical field. With the characteristic-size of subwavelength and multi-physics light field modulation, metasurfaces serve as SL projectors, offering significant advantages such as large FOV, compact size, and high-density information, which accelerate their application in 3D imaging.[24–28] Xiaoli Jing et al. projected a pseudo-random coding dot arrays into space using a metasurface, and utilized the local uniqueness of light dots to obtain 3D information, constituting an effective 3D imaging approach.[29] Kim and colleagues used a metasurface with a 180°FOV to generate dot arrays in the transmitted space, achieving 3D imaging of spatial objects.[30] However, the current research works are limited to the FOV within the transmission space, and the truly effective imaging range is smaller, making it still difficult to achieve a full-space effect. Additionally, these methods can only image a single picture and lack real-time capabilities owing to the computational complexity of the depth extraction algorithms, thus unable to reconstruct a complete 3D model of the target. The full-space FOV can contain 3D information of the entire 4π spatial range, which is conducive to a comprehensive perception of the surrounding environment, while the real-time capabilities allow for better interactive operations. Therefore, the pursuit of FOV maximization and real-time operation represents an extremely important direction for the future development of 3D reconstruction technology.

Based on this, we designed a metasurface to project approximately 20000 random dots into a 4π full-space. Combining binocular vision techniques with stereo matching algorithms, we compute the depth information of dynamic images in real-time to obtain the 3D point cloud. Then, we perform incremental point cloud reconstruction by stitching point cloud of consecutive frames based on cameras pose (**Figure 1a**). To verify the application prospects of the system in 3D scene reconstruction, we place the metasurface in the center of a real 3D space, aiming to uniformly fill the entire space with projected light dots (Figure 1b). Two cameras scan and capture the spatial scene, and a real-time mapping algorithm was used to synchronously construct the corresponding 3D point cloud (Figure 1c) and reconstruct its surface model (Figure 1d). The experiments results demonstrated that by using this metasurface, 3D models can be constructed for different scenes with a great range of space, and has excellent reconstruction quality (the overall reconstruction error is as low as 3.8‰) and high real-time performance (the processing speed

is at least 30 fps). Therefore, our approach offers a solution for simplifying the equipment, enhancing the FOV, and relaxing the applicable conditions for 3D reconstruction. It holds significant technological advantages in the future digitization process.

2. Results

2.1. Design and Characterization of the 4π Full-Space Metasurface

To effectively modulate light waves throughout the full-space, the diffractive structure must operate in both transmission and reflection modes simultaneously. Dielectric nanostructures can form scatterers that modulate the scattering of light waves utilizing their own Mie resonance characteristics.[31] By appropriately adjusting the wavelength of resonance peak, it is feasible to tailor the ratio of transmitted to reflected light energies, enabling the metasurface to simultaneously manipulate the light field in both transmission and reflection spaces. Based on this, the dielectric nanobrick is used as the unit diffraction cell for the full-space metasurface, as shown in **Figure 2a**. Its phase modulation of the incident light wave is given by geometric phase, i. e., $\varphi = \pm 2\alpha$, where α is the nanobrick orientation angle,[32,33] and the positive/negative sign depends on the rotation direction of the incident circularly polarized (CP) light. To ensure uniform energy distribution of the diffracted beam from the metasurface in diffraction space, the nanobrick structure is engineered to maintain equivalent modulation effects on the light wave across both the transmission mode and reflective mode. Here, we designed the metasurface with a working wavelength (λ) of 650 nm. For the nanobrick material, we chose amorphous silicon, which has high refractive index ($n = 3.227$) and low extinction coefficient ($k = 0.04$) at the working wavelength. The nanobrick arrays are fabricated on a silicon dioxide substrate. Through theoretical and simulation studies, we obtained an optimized nanobrick structure with a length (L) of 250 nm, width (W) of 120 nm, height (H) of 280 nm, and center spacing (C) of 300 nm. Figure 2b shows the efficiency of CP light emission from individual nanobrick. The efficiency of cross-polarized light modulation by this structure in the transmission and reflection spaces at the working wavelength is 30.23% and 27.90%, respectively, achieving an energy ratio close to 1:1. Additionally, the unwanted co-polarized light which contributes to the zero-order is compressed to be 3.31%, while the reflected CP light is 0.10% (see more detail in Figure S1, Supporting Information).

We designed a 4π full-space metasurface to project a dot arrays SL to facilitate the effective reconstruction of 3D model of scene within a 2 m side length range space. The periodic structure of the metasurface is closely related with the number of projected dots, as the distance between the metasurface and the object increases, the adjacent dots in the projection pattern separate further and single dot energy-density diminishes. Thus, analysis shows that reconstruction accuracy and range are well-maintained with ≈20 000 dots throughout the entire space. Since the nanobrick structure operates in both transmission and reflection modes, only 10 000 bright dots are required for the metasurface projection pattern. To avoid the distribution of the dots from being too dense, resulting in low energy per dot, or too sparse, leading to low imaging resolution, we optimized the design to have the bright dots occupy ≈25% of all controllable dots. Therefore, each period of the metasurface contains 200 × 200 nanobricks with a spacing of 300 nm, and a total of 5 × 5 periods are designed to reduce laser speckle. For $\lambda = 650$ nm, calculations show that the largest diffraction angle θ of the bright dots is 85.3° which allows for a FOV of 170.6° in a single transmission or reflection space. The overall diffraction dots almost cover the entire 4π full-space range. More details about the design are provided in Figure S2 (Supporting Information). The proposed metasurface is based on Fourier holography to project the SL into free space, and thus the design of the target diffraction pattern is crucial. As shown in Figure 2c, when a regularly shaped object is present in the target reconstruction scene, a uniform light dot arrays might merely avoid the object's edges, leading to insufficient contour information. In contrast, a random dot arrays, ensures that some light dots impinge on the edges of the object when there are enough dots, enabling the reconstruction result to encompass more comprehensive information about the scene. Therefore, the target diffraction pattern adopts the design of a random dot arrays. The designed metasurface, based on the geometric phase, possesses the Hermitian conjugate property in both transmission and reflection modes. Specifically, when the target diffraction pattern exhibits centro-symmetry, the sample produces the same far-field diffraction outcomes for both left-handed circularly polarized (LCP) light and right-handed circularly polarized (RCP) light. In a word, the target diffraction pattern we designed consists of a centrally symmetric array of 10 000 random dots, as depicted in Figure 2d. The phase distribution of the meta-hologram was restored using the Gerchberg-Saxton (G-S) algorithm. We fabricated a sample by standard electron-beam lithography. Figure 2e presents a photograph and a scanning electron microscope (SEM) image (partial view) of the metasurface. As a validation, we placed the sample in the center of a square box and illuminated it with a laser and the actual diffraction results are shown in Figure 2f, with an enlarged view of the local random dots in the bottom-right corner. It can be observed that our designed metasurface can project random light dots throughout the entire 4π space. Additionally, we characterized the relationship between the diffraction dots of the metasurface and the propagation angles in Figure S3 (Supporting Information).

2.2. Binocular Stereo Vision System for Extracting Depth Information of Objects in 3D Space

We constructed a binocular stereo vision system using two identical cameras to capture the reflected images of the random dot arrays projected by the metasurface, enabling the calculation of depth values of objects in 3D space, as shown in **Figure 3a**. The cameras were placed symmetrically with a certain deflection angle to expand the common FOV between the left and right images, thereby enhancing the range of single-shot depth imaging. Before the experiment, we calibrated the cameras using the Zhengyou Zhang calibration method.[34] Multiple images of a standard chessboard pattern were captured to obtain the intrinsic and extrinsic parameters of the cameras. These parameters are necessary for depth information extracting. The calibration process is described in Figure S4 (Supporting Information). The binocular stereo imaging system operates based on the triangle

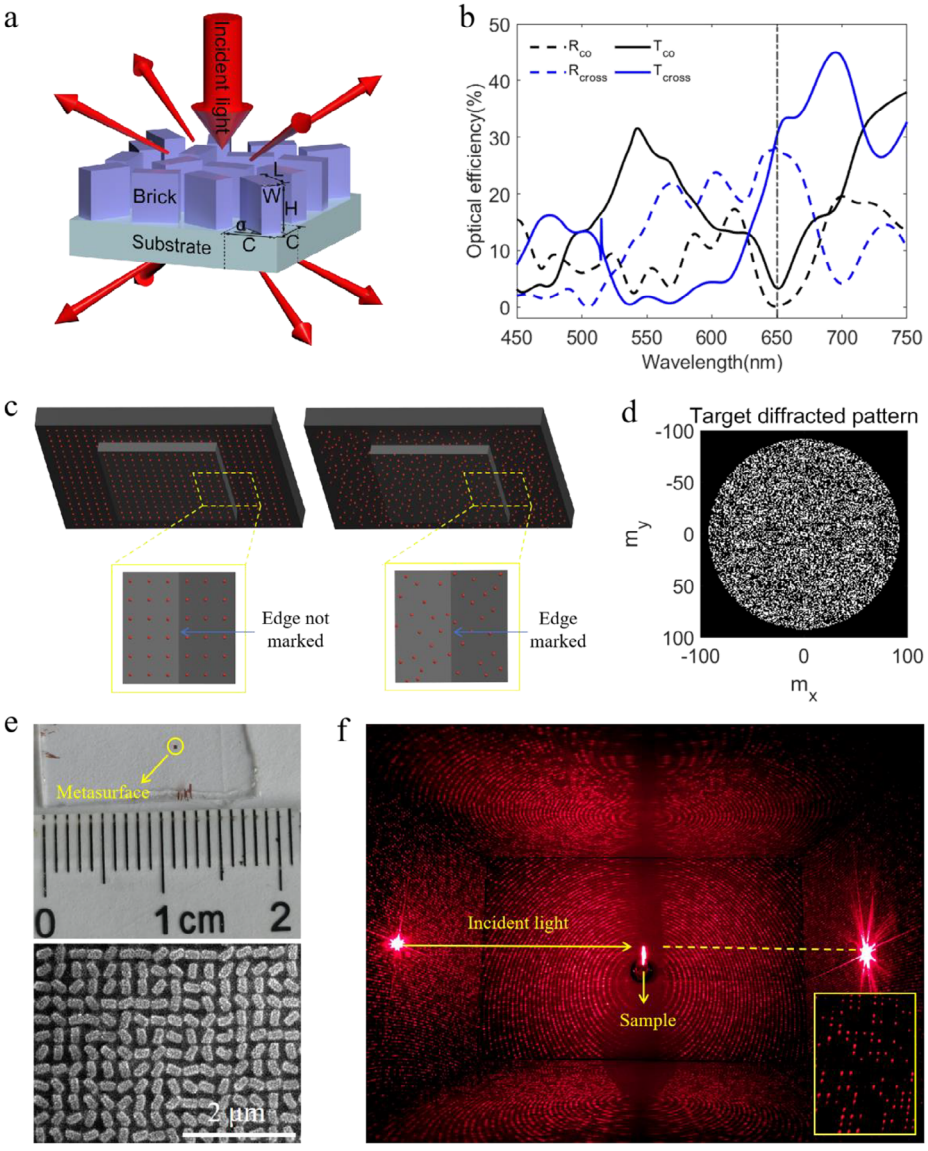

Figure 2. Design and experimental measurement of the full-space metasurface. a) Schematic illustration of the nanobricks spatial structure and light field modulation of the geometric phase metasurface. b) Efficiency of CP light emission from individual nanobrick, where T_{co} and R_{co} represent the efficiencies of co-polarized transmitted and reflected CP light, respectively, while T_{cross} and R_{cross} represent the efficiencies of cross-polarized transmitted and reflected CP light. c) Comparison of object marking effects between random dot arrays and uniform dot arrays, showing better marking performance of the random dot arrays for the edge of regular object. d) Design of the target diffraction pattern for the metasurface, containing 10 000 random bright dots within the circular region. e) The photograph and the partial SEM image of the metasurface with overall sizes of 300 μm × 300 μm. f) Experimental verification of the diffraction effect of the metasurface, demonstrating uniform filling of random light dots in both the transmission and reflection spaces, and the enlarged view of the random dots has been noted at the bottom-right corner.

similarity principle,[35] as schematically depicted in Figure 3b. There, O, M denote the center of the lens and the image plane of the camera respectively. The baseline is denoted as L, and P denotes a point in 3D space. α and β represent the angles between the camera and the baseline, as well as the angles between the camera and PO, respectively. x denotes the horizontal coordinate of P in the image coordinate system, and f is the focal length of the camera. The subscripts "l" and "r" correspond to the left and right cameras, respectively. Using geometric relations and trigonometric functions, a mathematical model for calculating the depth of object points in space has been established:

$$\begin{cases} Z = \frac{L}{cot(\alpha_l - \beta_l) + cot(\alpha_r - \beta_r)} \\ tan\beta_l = \frac{x_l}{f} \\ tan\beta_r = -\frac{x_r}{f} \end{cases} \quad (1)$$

when deflection angle $\alpha_l = \alpha_r = \pi/2$, the equation becomes:

$$Z = \frac{Lf}{x_l - x_r} \quad (2)$$

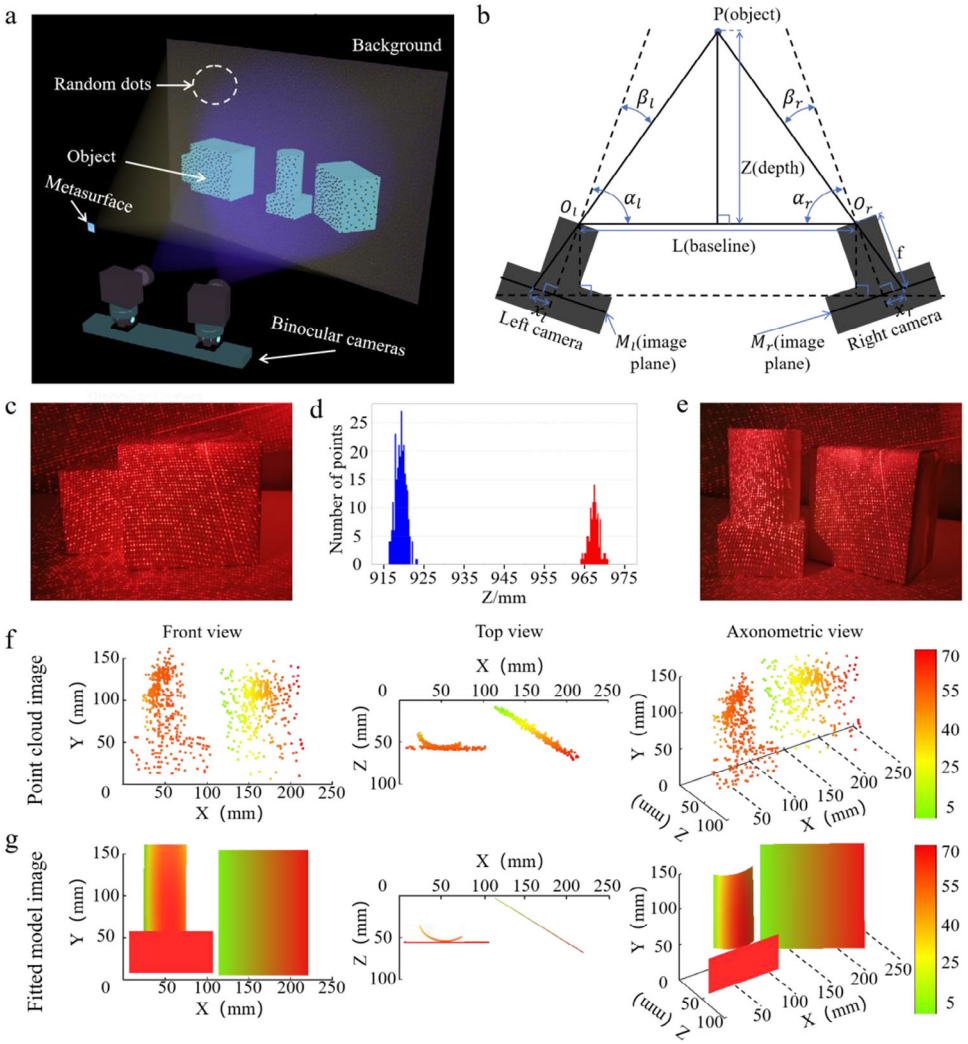

Figure 3. Depth extraction of space objects. a) Schematic diagram of the binocular imaging system for measure the depth values and surface morphology of objects. b) Schematic diagram of the triangle similarity principle. c) Depth measurement experiment with step distribution of the cardboard boxes. d) Depth distribution histogram corresponding to c), where the blue color represents the farther plane and the red color represents the nearer plane. e) Surface measurement experiment. f) Different perspectives on the spatial distribution of the 3D point cloud. g Surface model of the object obtained by fitting the 3D point cloud in f).

Here, $x_l - x_r$ represents the disparity, which means the difference in the projection of the same point in space on the image planes of two cameras. For detailed principles, please refer to Figure S5 (Supporting Information).

The primary method we employ to calculate the depth value involves identifying the corresponding point in the left and right images, and then deriving the depth information based on the relative positional relationship of the cameras and the image coordinates of the corresponding point. The whole process is segmented into epipolar rectification, feature extraction, stereo matching, and depth calculation. Our cameras are positioned non-parallel to each other, necessitating that depth cannot be directly calculated using equation 2. Using equation 1 for calculation would significantly increase computational load and greatly slow down the processing speed. Therefore, before the actual calculation, we employ the cameras' intrinsic and extrinsic parameters to perform distortion elimination and row alignment for the respective left and right images, termed epipolar rectification (see Figure S6, Supporting Information). After rectification, the left and right views only need to search for corresponding points along the same row, and equation 2 can be directly applied during depth calculation, greatly improving the efficiency of image processing. We extract the 2D coordinates of image points with obvious feature in the image, which can achieve sub-pixel precision, providing higher accuracy for subsequent matching. Then, we match the corresponding image points in the left and right views to calculate the disparity. To maintain matching quality while reduced computation demand, we use a Semi-Global Block Matching (SGBM) (Figure S7, Supporting Information) algorithm.[36–38] Ultimately, the derived disparity is used to calculate the depth value of each pixel point.

To validate the accuracy of extracting depth information from the constructed binocular stereo imaging system, we conducted a depth measurement experiment, as depicted in Figure 3c. Two

cardboard boxes were placed in space and projected with random dot arrays. The boxes had a step distribution from front to back, with a relative depth of 48 mm. After projecting with random dots by the metasurface, the surfaces of the cardboard boxes presented rich texture features. We captured the target images from the front using binocular cameras and performed depth calculation to obtain the depth data for the front face of the boxes. The depth values of all calculated points were plotted as a histogram, as shown in Figure 3d. The results show that the depths of the two planes are each uniformly distributed within a very narrow range. The average depth value of the farther plane is 967.280 mm, while the nearer plane is 919.184 mm, and the average distance between the two planes is 48.184 mm, the relative average error is 3.8% which proves the accuracy of the depth calculation for system. Figures S8 and S9 (Supporting Information) provide additional error data and spatial resolution used to characterize the accuracy. To further verify the system's performance in complex scenes, we carried out a surface measurement experiment as shown in Figure 3e. One cardboard box was rotated at a certain angle, and a cylinder was placed on the top surface of the other box. We extracted the depth information to observe the surface morphology. The depth values obtained from the surface measurement were visualized, and after denoising and filtering, they formed a 3D point cloud of the objects. Figure 3f shows the point cloud distribution from different perspectives. To visually represent the depth distribution and the surface reconstruction effect of the boxes, we fitted the obtained point cloud to generate a surface model, as shown in Figure 3g. The results of the two experiments demonstrate that the system is able to accurately measure the depth values of spatial objects and effectively reconstruct the complete morphological information of the scene. In addition, due to the application of epipolar rectification and semi-global feature matching approach in the depth recovery algorithm, our system achieves efficient single-shot image processing, enabling real-time processing speeds of at least 30 fps while ensuring accuracy. Combined with the large FOV of the metasurface, our system provides robust hardware and software support for real-time 3D scene reconstruction.

2.3. Demonstration of Real-Time 3D Scene Reconstruction with Large FOV

To demonstrate the accurate 3D reconstruction with a super-large FOV based on the metasurface projected dot arrays SL, we utilize the depth information from multiple frames of images (video) captured by the camera motion to stitch together a more extensive spatial point cloud by transforming the depth information according to the camera's pose. This allows us to establish a super-large FOV 3D model of the spatial scene. Here, we employ an algorithm called Simultaneous Localization and Mapping (SLAM) (Figure S10, Supporting Information).[39] We first establish two nodes for the left and right cameras through the Robot Operating System (ROS) system. These nodes publish topics to transmit image data and subscribe to topics to receive data, enabling real-time communication for two cameras and operating system, providing the live video feed as input to the algorithm. For each input frame, the depth information can be extracted using the binocular stereo imaging system. The algorithm utilizes the 3D coordinate changes of the same feature points between adjacent frames, combined with the rotation and translation matrices from the camera's intrinsic and extrinsic parameters, to calculate the coordinate transformation relationship. By solving for consecutive frames, we obtain the continuous motion trajectory of the camera during the scanning process. However, the cumulative error existing between frames can lead to significant deviations in trajectory under long time. To address this, we employ graph optimization in the back-end to incorporate information from all previous frames into the calculation. This approach distributes the errors evenly across each observation. Additionally, this method benefits from using sparse algebra for accelerated solving processes. To further ensure the accuracy of the camera poses, we use similarity detection to determine whether the captured images correspond to previously visited locations. If a match is found, the information that it is the same point is passed to the back-end optimization algorithm, which adjusts the trajectory based on this information. As a result, each frame's depth information corresponds to a set of cameras pose matrices. Through matrix transformations, we concatenate the depth values of all frames into the same coordinate system, resulting in a 3D point cloud of the scene captured by the camera during the scanning process. As the camera continues scanning, the calculation and point cloud generation proceed synchronously, leading to an increasing number of points in the point cloud. Hence, this reconstruction method is also known as incremental reconstruction. Using this characteristic, by simply scanning the space where the metasurface projects random dot arrays, we can obtain complete 3D point cloud data of the scene in space and reconstruct its 3D model.

As a validation, we set up four different scenes in a space measuring ≈1200 mm × 1000 mm × 600 mm. A laser was used to illuminate the metasurface placed in the center of the space, which generated a random dot arrays SL projection in both the reflection and transmission spaces. A binocular camera scanned the spatial scene, simultaneously acquire the 3D point cloud data of the scene, and measured the maximum viewing angle as large as 210° (Figure S11, Supporting Information). The real-time scanning and reconstruction process is shown in the Video S1 (Supporting Information). **Figure** 4a displays photographs of four different scenes. Scene 1 contains only the basic planar structure, with a right-angled corner in the metasurface reflection space and three in the transmission space; Scenes 2 and 3 each added a slant in the reflection or transmission spaces based on scene 1 (the slant is marked in the pictures with the yellow dashed box); Scene 4 contains two cubic boxes and a cylinder in both the transmission and the reflection spaces (shown by the yellow dashed outline in the picture). Each scene was scanned and reconstructed separately, resulting in their respective 3D point clouds. Figure 4b–d correspond to the front view, top view, and axonometric view of the point cloud distribution after denoising and filtering, respectively. To provide a more intuitive representation of the scenes, we constructed complete 3D models based on the point cloud results. Figure 4e shows the corresponding results with the perspective view taken from the camera position. Video S2 (Supporting Information) presents additional perspectives. The results demonstrate that the 3D models of each scene can be accurately reconstructed. Due to variations in the scanning range and partial occlusion of the experimental setup, there

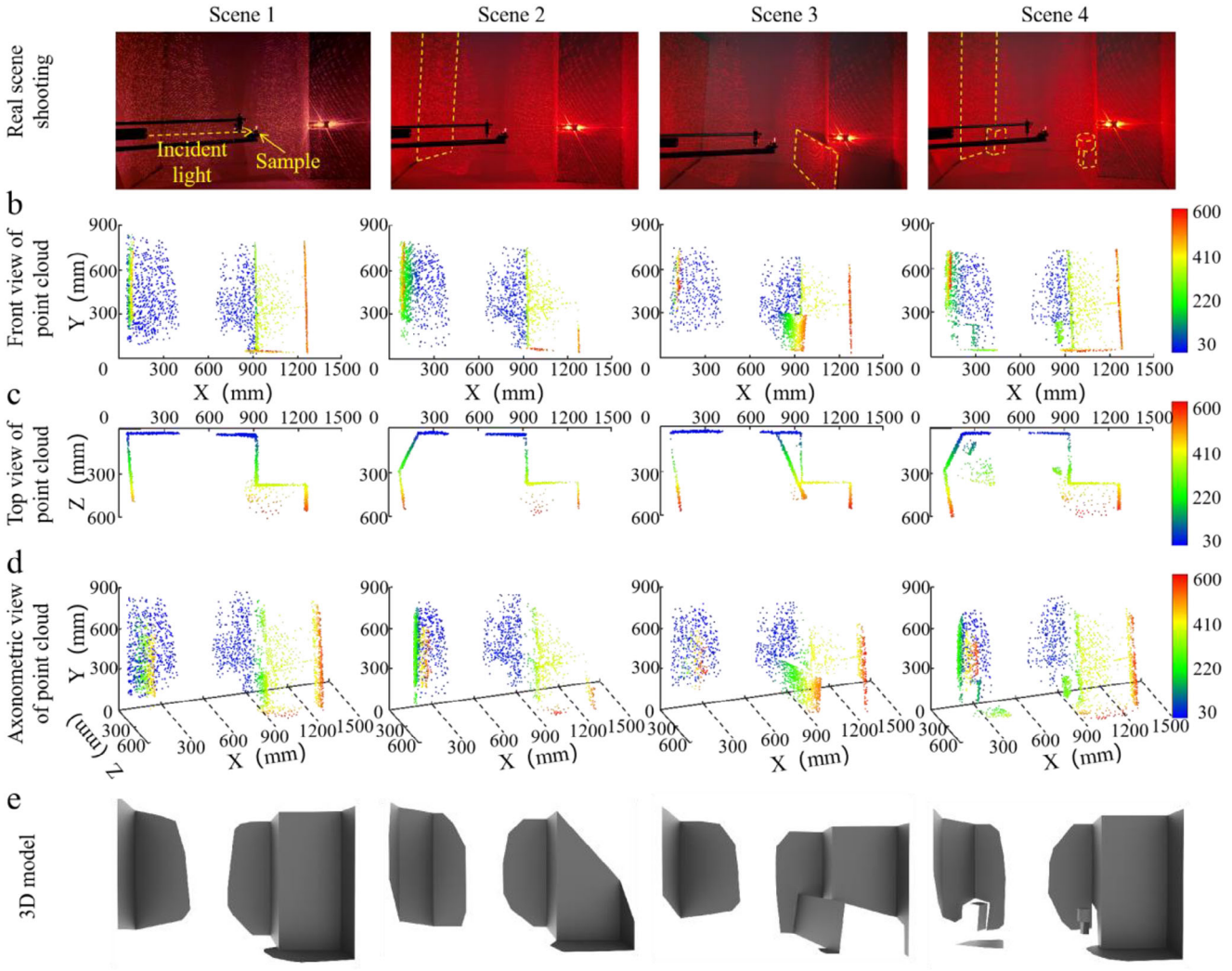

Figure 4. 3D reconstruction of four different spatial scenes. a) Photographs of the four different target scenes with projected random dot arrays. The yellow dashed boxes indicate the outlines of additional objects based on Scene 1. b–d) The incremental 3D point clouds distribution of four spatial scenes obtained from the scanning and calculation using the binocular imaging system. e Results of large FOV 3D modeling of point clouds of four scenes, demonstrating the perspective view with the real camera location as the viewpoint.

may be local missing areas in the reconstructed 3D models. However, in theory, as long as the scanning range is sufficiently large, our system can reconstruct the 3D model of the entire enclosed space.

3. Discussion

In our work, we leveraging the ultra-compact and wide-field of metasurfaces, the high-precision and real-time performance of depth extraction algorithms, as well as the unified reconstruction capabilities of point cloud simultaneous stitching algorithms, to construct a 3D reconstruction system with multiple advantages. This system is designed for applicability and practicality, rather than merely validating the metasurface's ability in 3D imaging. Compared to existing researches, our system fully exploits the 4π free space SL generated by the metasurface, extending the unified 3D reconstruction FOV to 210°. This indicates that the reconstruction module in our system is more compatible with the metasurface-based SL projector. Furthermore, our system achieved a relative error as low as 3.8%, allowing for the discernment of depth differences in 3D objects even in very small ranges. More importantly, we have, for the first time, achieved a real-time processing speed of at least 30 fps, which represents a significant breakthrough for 3D reconstruction based on metasurfaces. **Table 1** lists more information about the advantages of our work compared to related researches. In terms of experiments, we used a 10 mW laser to illuminate the metasurface, generating 20 000 random dots within a 1200 mm space range, resulting in high-resolution 3D reconstruction results. The quality of the reconstruction can be enhanced by employing finer metasurface design and fabrication techniques to reduce zero-order light. Using higher power lasers and increasing the number of dots can further expand the working range of reconstruction, while optimizing the point cloud stitching algorithm to reduce additional

Table 1. Comparison of advantages with related researches.

Research	Projector	FOV of SL	FOV of complete imaging	Imaging distortion	Relative depth accuracy	Result	Prepro-cessing	Real time
Dahnert et al.[1]	None	None	FOV of camera	None	No data	3D model	Yes	None
Choi et al.[16]	Meta	360°	<180°	Yes	35 mm (range: 2.5 m)	Depth map	Yes	None
Jing et al.[29]	Meta	90°	80°	None	0.2 mm (range: 0.3 m)	3D point cloud & depth map	Yes	None
Kim et al.[30]	Meta	180°	120°	None	40 mm (range: <1 m)	3D point cloud	None	None
Ni et al.[28]	Meta	120°	No imaging	–	–	–	–	–
Our	Meta	360°	210°	None	0.18 mm (range: ≈1 m)	3D point cloud and 3D model	None	Yes

sources of error associated with large-scale reconstruction. Alternatively, the system can be directly paired with other 3D measurement devices, such as LiDAR, to gather more information about the target. Given the flexibility of our method, the working wavelength can be tailored for other bands such as the infrared, as detailed in Figure S12 (Supporting Information). Then, by combining infrared cameras with color cameras, color 3D reconstruction can be effectively achieved, further enhancing the functionality of the system. Additionally, the metasurface can be integrated with vertical-cavity surface-emitting laser (VCSEL) to achieve a compact SL projector,[40] and the development of miniaturized cameras,[41] can lead to a lightweight SL detector, significantly reducing the overall system size. Overall, our research may contribute to further innovations and developments in advanced optical devices and 3D reconstruction technologies.

4. Conclusion

We designed and fabricated a metasurface with the equal transmissivity and reflectivity, enabling dot arrays SL projection filling the 4π full-space. Next, we construct a binocular vision imaging system that can reconstruct accurate 3D model of the scene in space in real-time based on stereo matching and point cloud transformation. In the experiment, we illuminated the designed metasurface to project 20 000 randomly distributed light dots into a cubic space. We then used a binocular camera to scan four different spatial scenes and simultaneously reconstruct their respective accurate 3D models, achieved an effective reconstruction FOV of 210° × 210°. The results demonstrate that the proposed metasurface-based dot arrays SL 3D reconstruction method exhibits super large imaging FOV and high real-time performance in real scene reconstruction, enabling instantly obtain 3D information of the entire scene and better perceive the environment in space. Compared to traditional laser point cloud-based 3D reconstruction techniques, our method achieves a significant leap in FOV coverage and demonstrates clear advantages in terms of system integration, paving the way for the development of cutting-edge fields such as real-time VR, holographic communication, and digital city construction. This indicates broader application scenarios and infinite possibilities for these fields in the future.

5. Experimental Section

Characterization of the Full-Space Property of the Metasurface: To demonstrate the projection capability of the designed metasurface with a 4π FOV, it was fixed it in the center of a cardboard box measuring 420 mm × 260 mm × 30 mm. Two small holes were opened on both sides of the box for beam entrance and exit. A beam with a wavelength of 650 nm was emitted from a laser, and vertically illuminate the metasurface. The laser beam was modulated by the metasurface, diffracting simultaneously into both the transmission and reflection spaces, forming a random dot arrays pattern. A high-resolution camera captured Figure 2f from a position ≈200 mm directly above the cardboard box.

Binocular Stereo Vision Imaging System Building: It was used two standard 5-megapixel industrial cameras assembled into a binocular stereo imaging system. The centers of the camera lenses were separated by $L = 200$ mm, with each camera deviating 5° toward the centerline. The focal length f for both cameras was 8 mm. A Python program was developed for camera calibration. The involved depth extraction algorithm, also implemented in Python, was responsible for calculating the depth values of object.

3D Point Cloud Data Processing: It was obtained 3D point cloud data for the corresponding targets in both the depth information extraction and real-time 3D scene reconstruction. Further processing was performed after removing interfering points using Cloud Compare. For the 3D point cloud obtained from depth information extraction, which had a relatively small number of points, a quadric surface fitting operation was performed. In contrast, for the richer point cloud obtained from scene reconstruction, a model based on the distribution of the point cloud was directly constructed.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research was supported by National Key Research and Development Program of China (2023YFB2804704), National Natural Science Foundation of China (12274047 and 12374278), Natural Science Foundation of Chongqing, China (Grant No. CSTB2023NSCQ-MSX0317) and Research Innovation Project for Postgraduate Students of Chongqing Jiaotong University (Grant No. 2023S0092).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

3D reconstruction, high-accurate, metasurface, real-time, structured light, super-large Field-of-View

Received: July 18, 2024
Revised: November 9, 2024
Published online:

[1] M. Dahnert, J. Hou, M. NieBner, A. Dai, presented at NeurIPS, 35th Conference on Neural Information Processing Systems, Virtual, December, 2021. arXiv: 2111.02444.
[2] P.-F. Zhang, X. Dai, J.-P. Ding, D.-W. Gong, S.-F. Wang, *Laser. Eng.* **2019**, *42*, 187.
[3] R. C. Shinde, S. S. Durbha, A. V. Potnis, *ISPRS JPRS* **2021**, *180*, 313.
[4] A. Maccarone, K. Drummond, A. McCarthy, U. K. Steinlehner, J. Tachella, D. A. Garcia, A. Pawlikowska, R. A. Lamb, R. K. Henderson, S. McLaaughlin, Y. Altmann, G. S. Buller, *Opt. Express* **2023**, *31*, 16690.
[5] Q. Zhang, H. Li, X. Wang, Q. Wang, *Int. J. Comput. Vis.* **2021**, *129*, 3006.
[6] X. Li, Z. Liu, Y. Cai, J. Yan, W. Wu, G. Guo, X. Shao, *Appl. Opt.* **2023**, *62*, 5627.
[7] M. Schaffer, M. Grosse, B. Harendt, R. Kowarschik, *Opt. Lett.* **2011**, *36*, 3097.
[8] P. Zhou, J. Zhu, W. Xiong, J. Zhang, *Appl. Opt.* **2021**, *60*, 5925.
[9] Z. Cai, G. Pedrini, W. Osten, X. Liu, X. Peng, *Opt. Lett.* **2020**, *45*, 3256.
[10] C. Tian, Z. An, L. Li, B. He, *Adv. Mater. Res.* **2010**, *910*, 145.
[11] J. Zhou, K. Qian, *Appl. Opt.* **2019**, *58*, A283.
[12] M. Servières, V. Renaudin, A. Dupuis, N. Antigny, *J. Sensors* **2021**, *2021*, 1.
[13] P. Zanuttigh, G. Marin, C. D. Mutto, F. Dominio, L. Minto, G. M. Cortelazzo, *Time-of-Flight and Structured Light Depth Cameras*, Springer, Cham, Berlin, Germany **2016**.
[14] Z. Li, Q. Dai, M. Q. Mehmood, G. Hu, B. Luk'Yanchuk, J. Tao, C. Hao, I. Kim, H. Jeong, G. Zheng, S. Yu, A. Alù, J. Rho, C.-W. Qiu, *Light: Sci. Appl.* **2018**, *7*, 63.
[15] Y. Wang, Q. Fan, T. Xu, *Opto-Electron. Adv.* **2021**, *4*, 200008.
[16] E. Choi, G. Kim, J. Yun, Y. Jeon, J. Rho, S.-H. Baek, *Nat. Photonics* **2024**.
[17] Y. Wang, C. Chen, S. Wu, X. Ye, S. Zhu, T. Li, *Chin. Opt. Lett.* **2024**, *22*, 053601.
[18] C. Chen, X. Xiao, X. Ye, J. Sun, J. Ji, R. Yu, W. Song, S. Zhu, T. Li, *Light: Sci. Appl.* **2023**, *12*, 288.
[19] S. Wang, S. Wen, Z.-L. Deng, X. Li, Y. Yang, *Phys. Rev. Lett.* **2023**, *130*, 123801.
[20] Z. Feng, T. Shi, G. Geng, J. Li, Z.-L. Deng, Y. Kivshar, X. Li, *eLight* **2023**, *3*, 21.
[21] L. Liu, W. Liu, F. Wang, X. Peng, D.-Y. Choi, H. Cheng, Y. Cai, S. Chen, *Light: Sci. Appl.* **2024**, *13*, 131.
[22] Q. Chen, G. Qu, J. Yin, *Nat. Nanotechnol.* **2024**, *19*, 1000.
[23] C. Liang, Q. Dai, T. Huang, H.-C. Liu, Z. Guan, R. Fu, J. Tao, Z. Li, S. Yu, G. Zheng, *J. Materiomics* **2024**, *10*, 811.
[24] X. Jing, Y. Li, J. Li, Y. Wang, L. Huang, *Nanophotonics* **2023**, *12*, 1923.
[25] D. B. Conkey, R. P. Trivedi, S. R. P. Pavani, I. I. Smalyukh, R. Piestun, *Opt. Express* **2011**, *19*, 3835.
[26] S. Tan, F. Yang, V. Boominathan, A. Veeraraghavan, G. V. Naik, *ACS Photonics* **2021**, *8*, 1421.
[27] W. Liu, D. Ma, Z. Li, H. Cheng, D. Y. Choi, J. Tian, S. Chen, *OpticaOptica* **2020**, *7*, 1706.
[28] Y. Ni, S. Chen, Y. Wang, Q. Tan, S. Xiao, Y. Yang, *Nano Lett.* **2020**, *20*, 6719.
[29] X. Jing, R. Zhao, X. Li, Q. Jiang, C. Li, G. Geng, J. Li, Y. Wang, L. Huang, *Nat. Commun.* **2022**, *13*, 7842.
[30] G. Kim, Y. Kim, J. Yun, S.-W. Moon, S. Kim, J. Kim, J. Park, T. Badloe, I. Kim, J. Rho, *Nat. Commun.* **2022**, *13*, 5920.
[31] C. F. Bohren, D. R. Huffman, *Absorption and scattering of light by small particles*, Wiley, Hoboken, NJ, USA **1983**.
[32] L. Huang, X. Chen, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, T. Zentgraf, S. Zhang, *Nano Lett.* **2012**, *12*, 5750.
[33] X. Chen, L. Huang, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, C.-W. Qiu, S. Zhang, T. Zentgraf, *Nat. Commun.* **2012**, *3*, 1198.
[34] Z. Zhang, *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330.
[35] D. Marr, T. Poggio, *Science* **1976**,*194*, 283.
[36] H. Hirschmüller, *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 328.
[37] S. Birchfield, C. Tomasi, *Int. J. Comput. Vision* **1999**, *35*, 269.
[38] H. Hirschmüller, presented at CVPR' 05, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recongnition, San Diego, CA, USA, June, 2005, https://doi.org/10.1109/CVPR.2005.56.
[39] R. Mur-Artal, J. D. Tardós, *IEEE T. Robot.* **2017**, *33*, 1255.
[40] Q. Wang, P. Ni, Y. Xie, Q. Kan, P. Chen, P. Fu, J. Deng, T. Jin, H. Chen, H.-W.-H. Lee, C. Xu, P. Genevet, *Laser Photonics Rev.* **2020**, *15*, 2000385.
[41] P. Evgeniy, W. Lukas, E.-G. Daniel, D.-J. Timothy, J.-D. Timothy, R. Ann, *Adv. Opt. Mater.* **2019**, *7*, 1900893.