ACS Photonics

pubs.acs.org/journal/apchd5

Article

High-Resolution, Wide-Field-of-View, and Real-Time 3D Imaging Based on Spatial-Temporal Speckle Projection Profilometry with a VCSEL Projector Array

Wei Yin,[⊥] Hang Zhao,[⊥] Yifan Ji, ZhongDa Deng, Ziheng Jin, Shijie Feng,* Xiaolei Zhang, Huai Wang, Qian Chen, and Chao Zuo*

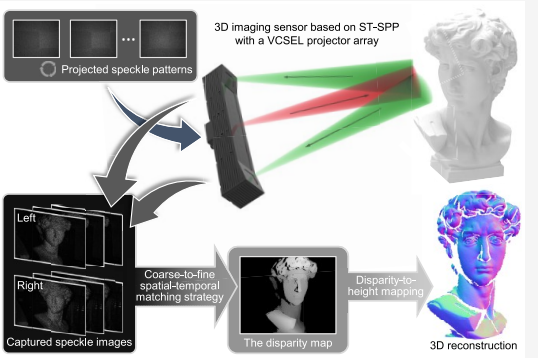
Cite This: *ACS Photonics* 2024, 11, 498−511

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

Downloaded via ZHEJIANG UNIV on February 23, 2024 at 08:50:46 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ABSTRACT: Structured light projection, especially fringe projection profilometry (FPP), is a mainstream high-precision noncontact 3D imaging technique for manufacturing, basic research, and engineering applications. However, FPP methods are severely limited by the complex structure and high cost of high-resolution spatial light modulation devices used for projecting high-quality fringe patterns, bringing tremendous challenges to miniaturized and low-cost 3D imaging applications. On the other hand, benefiting from advanced manufacturing processes, speckle projection profilometry (SPP), which utilizes highly integrated speckle projection devices based on a vertical-cavity surface-emitting laser (VCSEL), realizes fast 3D reconstruction by projecting only a single speckle pattern, opening up a new avenue for miniaturized depth sensing applications, such as scene reconstruction and face recognition. Nevertheless, SPP yields coarse 3D measurement results with low quality in terms of accuracy and resolution due to the poor performance of single-frame speckle matching and the complex reflection characteristics of the tested surfaces. In this paper, we present a high-resolution, wide-field-of-view, and real-time 3D imaging method using spatial-temporal speckle projection profilometry (ST-SPP), which integrates a set of VCSEL-based miniaturized speckle projectors to spatially and temporally encode the depth information on the measured scenes. A coarse-to-fine spatial-temporal matching strategy using digital image correlation is proposed to overcome the difficulty of applying speckle matching to complex surfaces, enabling high-precision and efficient subpixel disparity estimation. Experimental results prove that within the measurement range of 1 m (length) × 1.2 m (width) × 2 m (depth), ST-SPP achieves accurate and computationally efficient 3D imaging with a relative precision higher than 0.05%, exhibiting its advanced performance for long-range and real-time 3D reconstruction with fine details at 30 frames per second. The proposed ST-SPP is feasible for fast 3D modeling of dynamic scenes and large-scale objects with complex shapes, further enhancing the performance of optical metrology instruments based on SPP in terms of accuracy, resolution, measurement range, and portability.



KEYWORDS: *Speckle projection profilometry, 3D imaging, VCSEL, Stereo Digital Image Correlation, Speckle matching*

INTRODUCTION

Optical 3D measurement, as a promising metrology technique with the use of light as information carriers,[1,2] is extensively applied in intelligent manufacturing, scientific research, and medical diagnosis. With the advent of the laser[3] and charge-coupled device (CCD),[4] many optical metrology methods and instruments are employed in fast 3D modeling, robot navigation, microscopic 3D imaging,[5−7] and AR/VR, such as time-of-flight (TOF),[8−10] stereo vision,[11,12] and structured light projection.[13−20] TOF directly perceives the depth information on the target by recording the time difference from sending to receiving the modulated light signal, which is exclusively used for long-distance 3D sensing in driverless cars or visual navigation,[21] but its close-range measurement accuracy and resolution are relatively low. Stereo vision captures the intensity information on the scenes from two or more cameras and establishes the corresponding relationship between different perspectives by stereo matching,[22−26] realizing passive 3D imaging based on triangulation. The most compact stereo vision system consists of only two cameras with the advantages of simple hardware configuration and easy implementation. However, weakly textured regions hinder the search for matched points, limiting the accuracy and generalizability of stereo vision.

Received: September 19, 2023
Revised: December 22, 2023
Accepted: December 26, 2023
Published: January 31, 2024

ACS Publications

© 2024 American Chemical Society

Structured light projection, as an improved form of stereo vision, utilizes a light source generator to project patterns designed according to a series of coding strategies to enhance the surface features of objects, alleviating the mismatches in weakly textured regions. Common projected patterns mainly include fringe patterns and speckle patterns, which have been developed into two mainstream techniques: fringe projection profilometry (FPP)[27−31] and speckle projection profilometry (SPP).[32−35] For FPP, the projected fringe patterns are modulated by the objects and processed to recover the phase information related to the desired surfaces using various fringe analysis techniques.[36−41] However, these methods require performing the arctangent function for phase retrieval, resulting in the wrapped phases with 2π phase jumps. Phase unwrapping needs to be performed with additional patterns or more constraints to achieve absolute 3D measurements,[42−47] which limits its dynamic imaging performance.[48,49] In addition, as the core component of FPP systems, a digital light processing (DLP) projector is commonly used for projecting high-quality fringe patterns. However, FPP-based optical metrology instruments are severely limited by high cost, low optical power, and complex imaging systems of high-resolution DLP projectors, bringing tremendous challenges to miniaturized and long-range 3D imaging.

For SPP, a speckle pattern with a globally unique encoding distribution[50] is projected to assist in establishing accurate correspondence between stereo images, enabling single-shot 3D imaging. However, due to the complex reflection characteristics of tested surfaces and the perspective difference between multiple views, it is difficult to guarantee that every pixel in the whole measurement space has perfect global uniqueness by projecting one speckle pattern. To solve the common mismatching, some stereo matching algorithms, such as SGM[22,23] and PatchMatch,[26] are used to achieve robust 3D imaging by smoothing the matching results, but this leads to limited measurement accuracy and resolution. It is easy to understand that projecting multiple speckle patterns will improve the performance of 3D measurement because more spatial-temporal features can be exploited to enhance the global uniqueness of the measured scenes. Following this idea, Schaffer et al.[32,33] adopted an acousto-optical deflector to alter the laser beam direction, thereby changing the illuminated area of a ground-glass diffuser for achieving fast scanning of speckle patterns. Then, a temporal speckle correlation based on normalized cross-correlation (NCC) is exploited to enhance the measurement accuracy using 15 speckle images. However, this technique requires careful adjustment of the axial diffuser position in focus to fine-tune the speckle size, resulting in the quality deterioration of speckle patterns as the measurement range increases. Zhou et al.[51] proposed a speckle projector based on the binary mask wheel for generating real-time rotating speckle patterns. To ensure a low correlation between two adjacent speckle frames, the switching period of the speckle patterns takes 11 ms, which limits the capability of high-speed projection.

In addition, it seems at best naïve that existing spatial-temporal speckle matching methods directly extend local stereo matching algorithms to the spatial-temporal version, as rich spatial-temporal speckle information has not yet been fully exploited in current practice. Zhu et al.[52,53] adopted ZNCC-based spatial-temporal block matching for reliable integer-pixel disparity estimation, which is further refined by a five-point quadratic curve fitting for subpixel disparity optimization. The fundamental assumption made by block matching is that all of the pixels in the matching window have similar disparities. As a consequence, this assumption does not hold at disparity discontinuities, causing matching results with the edge-fattening issue in object boundaries and thin structures.[54,55] To overcome this challenge, the slanted support window proposed in PatchMatch,[26] as one of the state-of-the-art stereo matching algorithms, breaks the shackles of traditional fixed-window local matching and achieves 3D imaging with excellent performance. However, it requires performing multiple iterations for the local optimum of the slanted planes by the spatial, view, and temporal propagation, at the cost of expensive computational overhead and disparity smoothing.[56] Tang et al. proposed a spatial-temporal correlation method based on Newton−Raphson iteration for subpixel interpolation to improve the measurement precision, but cannot reconstruct complex surfaces with large changes in curvature.[57] On the contrary, temporal matching methods enable theoretically higher measurement accuracy and resolution than spatial-temporal correlation, but it requires at least 15 projected patterns and complex bicubic interpolation to enhance subpixel matching precision, which is not suitable for accurate and efficient 3D reconstruction of dynamic scenes.[30,58,59]

At present, existing spatial-temporal speckle generation methods project high-quality speckle patterns within the reduced measurement range through self-developed projection systems with complex structures, bringing tremendous challenges to highly integrated, cost-effective, and long-range 3D imaging instruments. Moreover, due to the inherent defect of block matching theoretically, simple but computationally expensive spatial-temporal speckle correlation algorithms still have a certain gap with FPP methods in terms of measurement accuracy and resolution, which cannot meet the requirements of precise measurement in industrial scenarios. In this paper, we present a high-resolution, wide-field-of-view, and real-time 3D imaging method using spatial-temporal speckle projection profilometry (ST-SPP) with a vertical-cavity surface-emitting laser (VCSEL) projector array. VCSEL is commonly adopted as the core light source of speckle projection modules in the industry world to simplify the manufacturing process of SPP systems,[60,61] which has the advantages of high integration, high optical power, and long probe distance. The proposed ST-SPP, by integrating a set of VCSEL-based speckle projectors, implements efficient spatial-temporal encoding for the depth information on the measured scenes at a pattern switching period of less than 1 ms. In addition, we prove that the slanted window model in PatchMatch can be expressed as a special form of the first-order shape function in the digital image correlation (DIC). Furthermore, a coarse-to-fine spatial-temporal matching strategy using DIC is proposed to overcome the difficulty of applying speckle matching to complex surfaces, enabling efficient and high-precision subpixel disparity estimation. Experimental results demonstrate that ST-SPP can achieve accurate and real-time 3D imaging within the wide-field-of-view and long-range measurement space of 1 m (length) × 1.2 m (width) × 2 m (depth).

METHOD

Long-range and wide-field-of-view 3D imaging sensor based on ST-SPP with a VCSEL projector array. Common 3D imaging sensors based on speckle projection adopt a single VCSEL-based speckle projection module to project the speckle pattern with a fixed spatial distribution into the

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

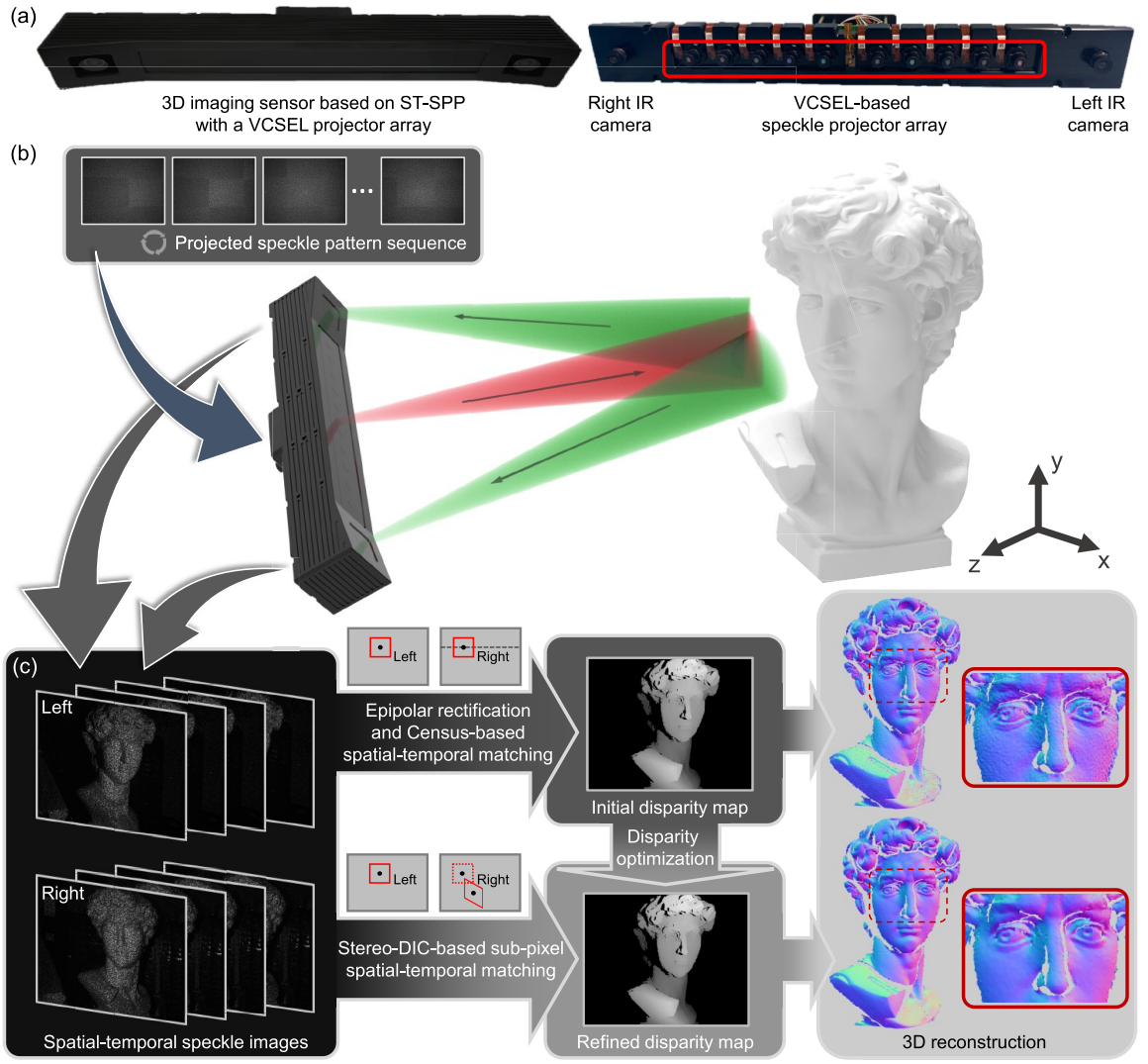


Figure 1. Overview of ST-SPP framework. (a) Photographs of a 3D imaging sensor based on ST-SPP with a VCSEL projector array and its internal structure, (b) spatial-temporal speckle projection and acquisition procedure, and (c) flowchart of data processing in ST-SPP for 3D reconstruction of the David model using a coarse-to-fine spatial-temporal matching strategy.

measured scenes. The global correspondences between speckle images captured by left and right cameras can be established using stereo matching, thereby achieving single-shot 3D reconstruction. However, due to the complex reflection characteristics of the measured surfaces and the perspective differences between stereo cameras, it is still difficult to encode a globally unique feature for every pixel in the whole measurement space by projecting a single speckle pattern, resulting in its limited measurement accuracy and resolution.

In order to achieve high-resolution, wide-field-of-view, and real-time 3D measurement, as shown in Figure 1a, the proposed ST-SPP-based 3D imaging sensor with a wide baseline integrates 10 speckle projection modules distributed at intervals in the horizontal direction to realize spatial-temporal speckle projection. For a VCSEL-based miniaturized speckle projector, dense and high-quality speckle projection is achieved by customizing the projection pattern of VCSEL and using the beam copy function of Diffractive Optical Elements (DOE; see Supporting Information for detailed analysis). In terms of hardware, our 3D sensor is configured with a 32-bit ARM microcontroller (GigaDevice, GD32F330) to perform the projection and acquisition of speckle patterns simultaneously. The GPIO port of the ARM microcontroller is adopted to control the turn-on and turn-off of the MOSFET at a driving voltage of 2.2 V, thus realizing rapid spatial-temporal speckle projection at a pattern switching period of less than 1 ms, while enabling the efficient and unique spatial-temporal encoding for the depth information on the measured scenes.

In addition, two infrared CMOS cameras (IR camera, Smartsens SC130GS) are operated at the SXGA resolution (1280 × 1024 pixels) to capture stereo speckle images synchronously at a speed of 30 Hz in Figure 1b. Since spatial-temporal speckle matching is flexible in the length of the image sequence to be processed, a sliding temporal window technique can be used to generate a new 3D reconstruction frame for every new speckle image pair captured at 30 FPS. The distance between left and right IR cameras is set to 0.4 m, which is also the baseline of 3D sensors. To maximize the valid measurement area, there is a certain intersection angle (about 10°) between two cameras and the z axis to guarantee long-range and wide-field-of-view 3D imaging within the measurement range of 1 m (length) × 1.2 m (width) × 2 m (depth). Moreover, a high-power infrared floodlight is placed in the center of the developed

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

3D sensor to provide uniform illumination for the collection of scene textures and system calibration.

Coarse-to-Fine Spatial-Temporal Matching Strategy Using Digital Image Correlation. For ST-SPP, the aim of spatial-temporal matching is to estimate a dense and high-precision disparity map from stereo speckle image sequences for 3D reconstruction of the target scene. The proposed coarse-to-fine spatial-temporal matching strategy consists of two steps: census-based spatial-temporal matching and Stereo-DIC-based subpixel spatial-temporal matching, as shown in Figure 1c. It is worth noting that before stereo matching, epipolar rectification is first executed to align the epipolar lines of speckle images, simplifying the two-dimensional search problem to a one-dimensional matching problem.

Census-Based Spatial-Temporal Matching. Specifically, a reliable and smooth initial disparity map of the tested scene is first obtained by using spatial-temporal matching based on a census transform. In our method, census transform with a local spatial-temporal window is adopted to extract the feature vectors for every pixel in speckle image sequences, which can be described as

$$C(x, y) = \bigotimes_{t=1}^{N} \bigotimes_{i=-R}^{R} \bigotimes_{j=-R}^{R} T(I(x+i, y+j, t), I_m(x, y)) \tag{1}$$

$$T(a, b) = \begin{cases} 0, & a \le b \\ 1, & a > b \end{cases} \tag{2}$$

$$I_m(x, y) = \frac{\sum_{t=1}^{N} \sum_{i=-R}^{R} \sum_{j=-R}^{R} I(x+i, y+j, t)}{N \times (2R+1)^2} \tag{3}$$

where $C(x, y)$ is the feature vector of the central pixel (x, y), $\otimes$ represents a bit-wise concatenation operator, and $I_m(x, y)$ is the mean intensity of the local spatial-temporal window. R and N are the spatial window radius and temporal length of spatial-temporal matching.

Based on the disparity range $[D_{\min}, D_{\max}]$ determined by our system calibration parameters and the preset measurement range, in Figure 1c, the matching cost $\mathrm{Cost}(x, y, d)$ can be obtained to estimate the similarity between each pixel in the left image and all candidates of the right image by calculating the Hamming distance of their feature vectors, which is defined as

$$\mathrm{cost}(x, y, d) = \frac{\mathrm{BC}(C_{\mathrm{L}}(x, y) \oplus C_{\mathrm{R}}(x-d, y))}{N \times (2R+1)^2} \tag{4}$$

where $\oplus$ is an XOR operation and $\mathrm{BC}(\cdot)$ is used to count the number of "1" in the XOR result. Then, the initial integer-pixel disparity map can be selected as the index of the minimum cost in $\mathrm{Cost}(x, y, d)$ through Winner Take All (WTA):

$$D(x, y) = \underset{d}{\operatorname{argmin}}\ \mathrm{cost}(x, y, d) \tag{5}$$

Further, subpixel disparity estimation is implemented by fitting a parabola using neighboring costs:

$$D^{\mathrm{sub}} = D - \frac{\mathrm{cost}(D+1) - \mathrm{cost}(D-1)}{2\,\mathrm{cost}(D+1) + 2\,\mathrm{cost}(D-1) - 4\,\mathrm{cost}(D)} \tag{6}$$

In addition, how to select the appropriate window radius, R, for speckle matching is our primary focus. It is obvious that the smaller R can speed up the calculation efficiency but only provides a coarse correlation result. On the contrary, the bigger R outputs reliable corresponding points at the cost of expensive computational overhead. In order to juggle the matching accuracy, spatial resolution, and computational overhead as much as possible, the smaller R should be adopted to obtain a coarse disparity map in a short time, and then a few points with low reliability (such as $\mathrm{Cost}(D) \le 0.8$) can be matched correctly using the larger window radius.

Stereo-DIC-Based Subpixel Spatial-Temporal Matching. The fundamental assumption of local matching is that all of the pixels in the matching window have similar disparities. However, this assumption does not hold at disparity discontinuities, leading to final 3D reconstruction results with the edge-fattening issue in object boundaries and thin structures. In order to overcome the low accuracy of traditional local spatial-temporal matching, a Stereo-DIC-based subpixel spatial-temporal matching is proposed to recover the fine profiles of the measured objects, thereby alleviating the disparity smoothing of local matching and further enhancing the performance of spatial-temporal matching in terms of both measurement accuracy and spatial resolution.

Both Stereo Digital Image Correlation (Stereo-DIC)[62,63] and SPP can be applied to measure 3D profiles of the tested surfaces. Similar to local speckle matching, Stereo-DIC matches local windows in the same way and then estimates the similarity between the reference and target window. However, Stereo-DIC firmly believes that the corresponding window on the target image may no longer be the same rectangle as the reference window due to the perspective differences between stereo cameras. In Stereo-DIC, the second-order shape function $\mathbf{W}(\xi; \mathbf{T})$ is introduced to describe a perspective transformation between the left reference subset centered on point $\mathbf{p}$ and the right target subset centered on point $\mathbf{q}$[64] as shown in Figure 1c:

$$\sum_{\xi} I_{\mathrm{L}}(\mathbf{p}+\xi) = \sum_{\xi} I_{\mathrm{R}}(\mathbf{p} + \mathbf{W}(\xi; \mathbf{T})) \tag{7}$$

$$\mathbf{W}(\xi; \mathbf{T}) = \begin{bmatrix} x + u + u_x x + u_y y + u_{xy} xy + \frac{u_{xx}}{2} x^2 + \frac{u_{yy}}{2} y^2 \\ y + v + v_x x + v_y y + v_{xy} xy + \frac{v_{xx}}{2} x^2 + \frac{v_{yy}}{2} y^2 \end{bmatrix} \tag{8}$$

$$\begin{cases} x_{\mathbf{p}} + u = x_{\mathbf{q}} \\ y_{\mathbf{p}} + v = y_{\mathbf{q}} \end{cases} \tag{9}$$

where $\xi = [x, y]^T$ denotes the local coordinate in the subset and $\mathbf{T}$ represents the deformation parameter vector of the target window, i.e., $[u, u_x, u_y, u_{xx}, u_{yy}, u_{xy}, v, v_x, v_y, v_{xx}, v_{yy}, v_{xy}]^T$. In addition, we prove that the slanted window model in PatchMatch can be expressed as a special form of the first-order shape function in Stereo-DIC (see Supporting Information for detailed analysis). Therefore, Stereo-DIC with the second-order shape function is expected to overcome the matching errors caused by the perspective differences between stereo cameras, maintaining a good trade-off between the computational efficiency and spatial resolution.

For Stereo-DIC-based subpixel optimization, a practical and robust spatial-temporal ZNSSD criterion is combined with a second-order shape function based on perspective trans-

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

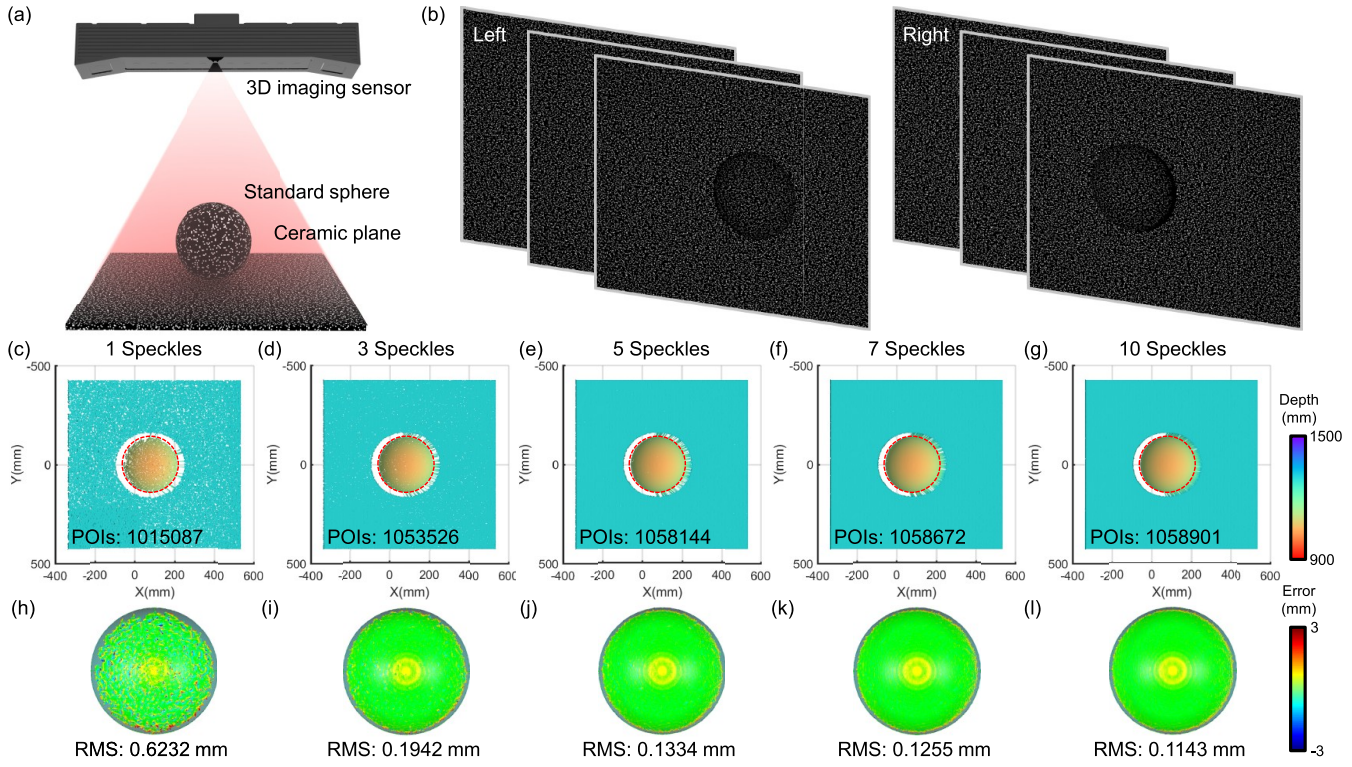


Figure 2. Simulation results of ST-SPP using different speckle patterns. (a) The simulated scene consisting of a ceramic plane and a standard sphere. (b) Stereo speckle image sequence captured by virtual left and right cameras. (c–l) 3D reconstruction results and the corresponding measurement errors using different speckle patterns.

formation to quantitatively evaluate the similarity between the left and right spatial-temporal subsets:

$$C_{\mathrm{ZNSSD}}(\Delta\mathbf{T}) = \sum_{t=1}^{N}\sum_{\xi}\left[\frac{I_{\mathrm{L}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\Delta\mathbf{T}))-\overline{I}_{\mathrm{L}}}{\Delta I_{\mathrm{L}}} - \frac{I_{\mathrm{R}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\mathbf{T}))-\overline{I}_{\mathrm{R}}}{\Delta I_{\mathrm{R}}}\right]^{2} \tag{10}$$

$$\Delta I_{\mathrm{L}} = \sqrt{\sum_{t=1}^{N}\sum_{\xi}[I_{\mathrm{L}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\Delta\mathbf{T}))-\overline{I}_{\mathrm{L}}]^{2}} \tag{11}$$

$$\Delta I_{\mathrm{R}} = \sqrt{\sum_{t=1}^{N}\sum_{\xi}[I_{\mathrm{R}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\mathbf{T}))-\overline{I}_{\mathrm{R}}]^{2}} \tag{12}$$

where $\mathbf{x}$ denotes the coordinate of the point to be matched in the left image, $\overline{I}_{\mathrm{L}}$ and $\overline{I}_{\mathrm{R}}$ represent the mean intensity values of the left and right spatial-temporal subsets. $\mathbf{W}(\xi;\Delta\mathbf{T})$ is the incremental shape function exerted on the left subset, and $\Delta\mathbf{T}$ represents the incremental deformation parameter vector:

$$\mathbf{W}(\xi;\Delta\mathbf{T}) = \begin{bmatrix} x+\Delta u+\Delta u_x x+\Delta u_y y+\Delta u_{xy}xy+\frac{\Delta u_{xx}}{2}x^2+\frac{\Delta u_{yy}}{2}y^2 \\ y+\Delta v+\Delta v_x x+\Delta v_y y+\Delta v_{xy}xy+\frac{\Delta v_{xx}}{2}x^2+\frac{\Delta v_{yy}}{2}y^2 \end{bmatrix} \tag{13}$$

To minimize the ZNSSD coefficient, a first-order Taylor expansion with respect to $\Delta\mathbf{T}$ can be performed:[63]

$$C_{\mathrm{ZNSSD}}(\Delta\mathbf{T}) = \sum_{t=1}^{N}\sum_{\xi}\left[\frac{I_{\mathrm{L}}^{t}(\mathbf{x}+\xi)+\nabla I_{\mathrm{L}}^{t}\frac{\partial\mathbf{W}}{\partial\mathbf{T}}\Delta\mathbf{T}-\overline{I}_{\mathrm{L}}}{\Delta I_{\mathrm{L}}} - \frac{I_{\mathrm{R}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\mathbf{T}))-\overline{I}_{\mathrm{R}}}{\Delta I_{\mathrm{R}}}\right]^{2} \tag{14}$$

where $\nabla I_{\mathrm{L}}^{t}$ is the gradients of the reference subset. A least-squares solution can be provided by minimizing the $C_{\mathrm{ZNSSD}}(\Delta\mathbf{T})$ with respect to $\Delta\mathbf{T}$, i.e., ${}^{\partial C_{\mathrm{ZNSSD}}(\Delta\mathbf{T})}/_{\partial(\Delta\mathbf{T})}$:[65]

$$\Delta\mathbf{T} = -\mathbf{H}^{-1}\sum_{t=1}^{N}\sum_{\xi}\left\{\left(\nabla I_{\mathrm{L}}^{t}\frac{\partial\mathbf{W}}{\partial\mathbf{T}}\right)^{T}\left[I_{\mathrm{L}}^{t}(\mathbf{x}+\xi)-\overline{I}_{\mathrm{L}} - \frac{\Delta I_{\mathrm{L}}}{\Delta I_{\mathrm{R}}}(I_{\mathrm{R}}^{t}(\mathbf{x}+\mathbf{W}(\xi;\mathbf{T}))-\overline{I}_{\mathrm{R}})\right]\right\} \tag{15}$$

where $\mathbf{H}$ is the 6×6 Hessian Matrix given as

$$\mathbf{H} = \sum_{t=1}^{N}\sum_{\xi}\left[\left(\nabla I_{\mathrm{L}}^{t}\frac{\partial\mathbf{W}}{\partial\mathbf{T}}\right)^{T}\times\left(\nabla I_{\mathrm{L}}^{t}\frac{\partial\mathbf{W}}{\partial\mathbf{T}}\right)\right] \tag{16}$$

Based on the calculated deformation parameter vector $\Delta\mathbf{T}$, the incremental shape function $\mathbf{W}(\xi;\Delta\mathbf{T})$ of the reference subset can be utilized to update the second-order shape function $\mathbf{W}(\xi;\mathbf{T})$ of the target subset:[65]

$$\mathbf{W}(\xi;\mathbf{T}) \leftarrow \mathbf{W}(\xi;\mathbf{T})\mathbf{W}^{-1}(\xi;\Delta\mathbf{T}) \tag{17}$$

It is noted that the initial guess of $\mathbf{T}$ can be estimated using least-squares-based surface fitting. With matching results obtained by census-based spatial-temporal matching according to eq 6, eq 7 can be rewritten based on the surface fitting:[66]

$$\mathbf{p}+\mathbf{W}(\xi;\mathbf{T}) = \mathbf{p}+\xi-D^{\mathrm{sub}}(\mathbf{p}+\xi) \tag{18}$$

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

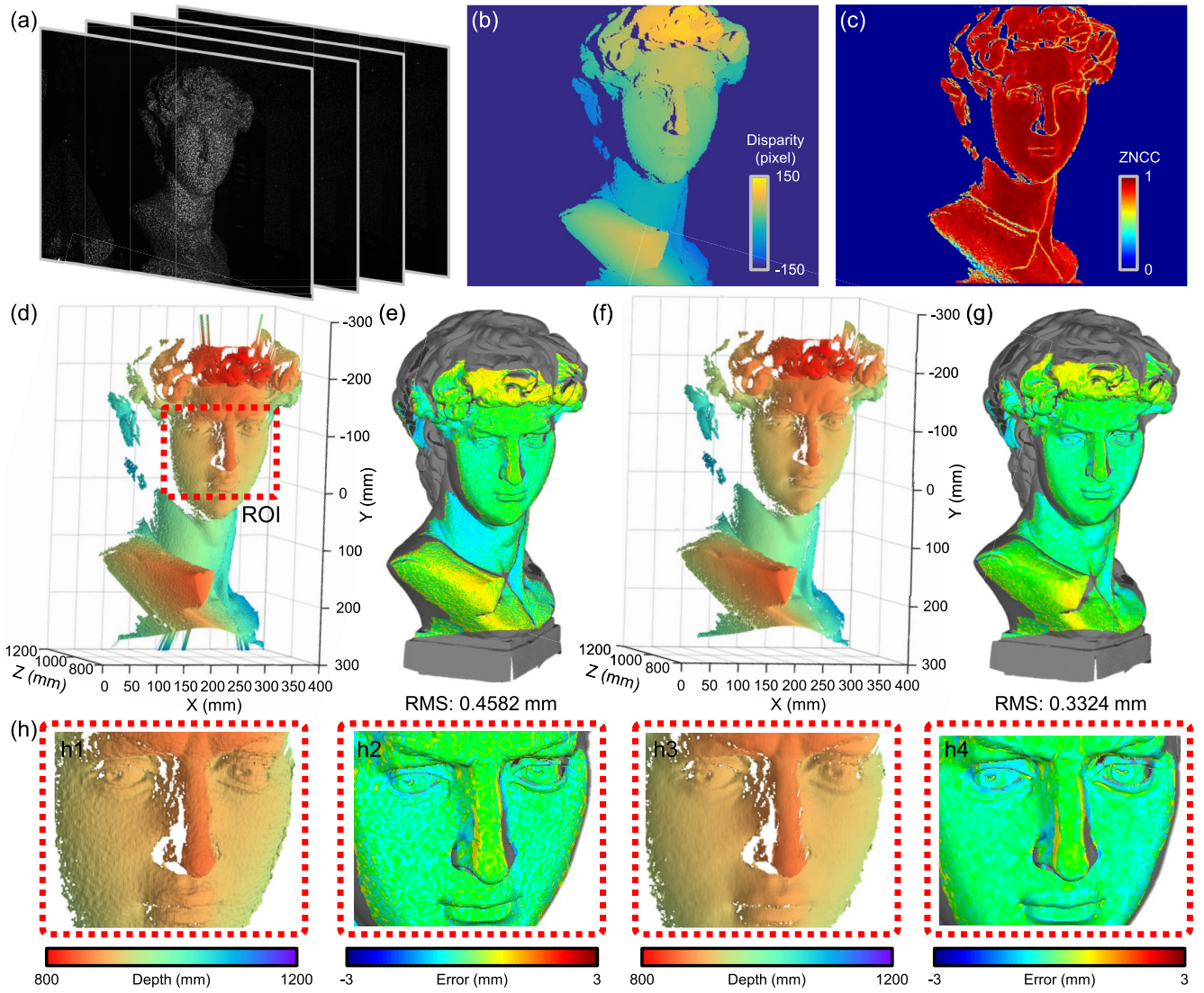


Figure 3. Large-scale 3D reconstruction of the David model. (a) Spatial-temporal speckle images. (b,c) The disparity map and ZNCC-based similarity output by ST-SPP. (d,e) 3D reconstruction results and the measurement errors by subpixel disparity estimation based on parabola fitting. (f,g) 3D reconstruction results and the measurement errors by Stereo-DIC-based subpixel spatial-temporal matching. (h) The magnified views of d−g.

$$\begin{cases} S_x^2 = \sum_{\xi} (a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 + D^{\text{sub}})^2 \\ S_y^2 = \sum_{\xi} (b_1x^2 + b_2y^2 + b_3xy + b_4x + b_5y + b_6)^2 \end{cases} \tag{19}$$

where $(a_1, a_2, a_3, a_4, a_5, a_6)$ and $(b_1, b_2, b_3, b_4, b_5, b_6)$ are set as $(u_{xx}/2, u_{yy}/2, u_{xy}, u_x, u_y, u)$ and $(v_{xx}/2, v_{yy}/2, v_{xy}, v_x, v_y, v)$. Solving eq 19 is a least-squares minimization problem for obtaining initial deformation parameters. In addition, to further speed up the initial estimation, it is optional to estimate only six parameters without considering the second-order coefficients. Finally, eq 15 is calculated iteratively until the predetermined convergence conditions are met. The convergence conditions are set to ensure that variations in the 2-norm of the incremental deformation parameter vector $\Delta\mathbf{T}$ are equal to or less than 0.001, i.e., $\|\Delta\mathbf{T}\|_2 \leq 0.001$.

3D Imaging Performance Analysis of ST-SPP Using Different Speckle Patterns. To evaluate the 3D imaging performance of ST-SPP using different speckle patterns, we built a virtual ST-SPP system using Blender to measure the simulated scene consisting of a ceramic plane and a standard sphere in Figure 2a. The speckle light source moves along the horizontal direction to simulate spatial-temporal speckle projection. Stereo speckle image sequences captured by virtual left and right cameras are shown in Figure 2b. Due to the fundamental assumption of similar disparities in block matching, measuring the objects with ridged, complex, or curved surfaces is a challenging task for SPP. For the obtained 3D reconstruction results using different speckle patterns in Figure 2c−g, it can be found that projecting a single-frame speckle image can only obtain rough measurement results with 1 015 087 valid points of interest (POIs), which have a large number of mismatching errors with a root-mean-square (RMS) of 0.6232 mm for measuring the standard sphere. As the projected speckle patterns are increased to three images, the mismatching errors in the 3D measurement results gradually decrease, the number of valid points improves to 1 053 526 POIs, and its measurement accuracy of the sphere surface is greatly enhanced from 0.6232 to 0.1942 mm by sphere fitting in Figure 2h,i. As the spatial-temporal speckle pattern sequence is expanded to 5, 7, and 10 frames, more accurate 3D reconstruction results with higher

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

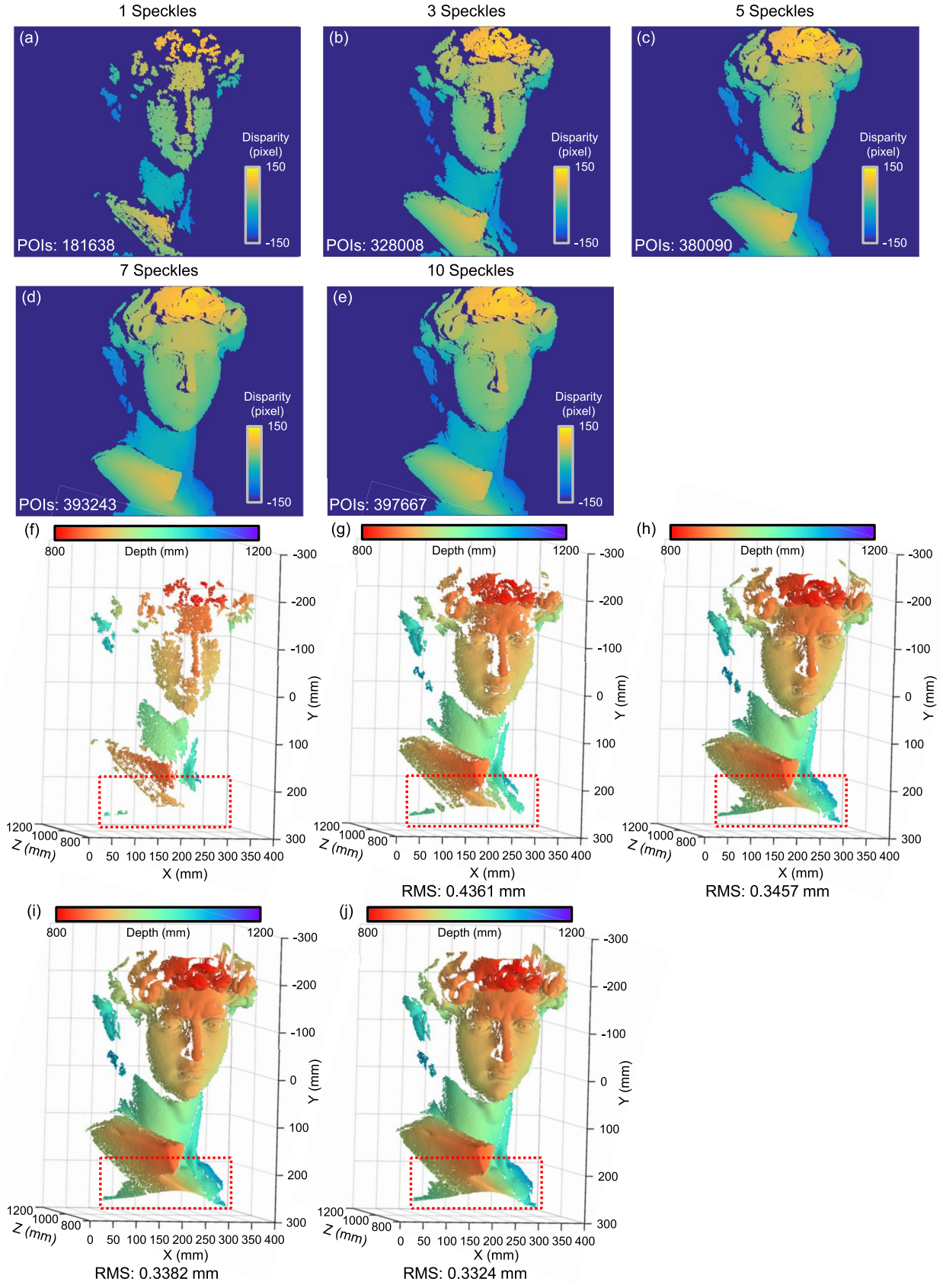


Figure 4. 3D reconstruction of the David model using different speckle patterns. (a–j) The disparity maps output and 3D reconstruction results by ST-SPP using one, three, five, seven, and 10 speckle patterns.

completeness can be obtained, verifying the reliability of the proposed ST-SPP for 3D imaging of curved surfaces, as shown in Figure 2c–l. Simulation experiments demonstrate that more robust 3D measurements with the highest accuracy can be

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498–511

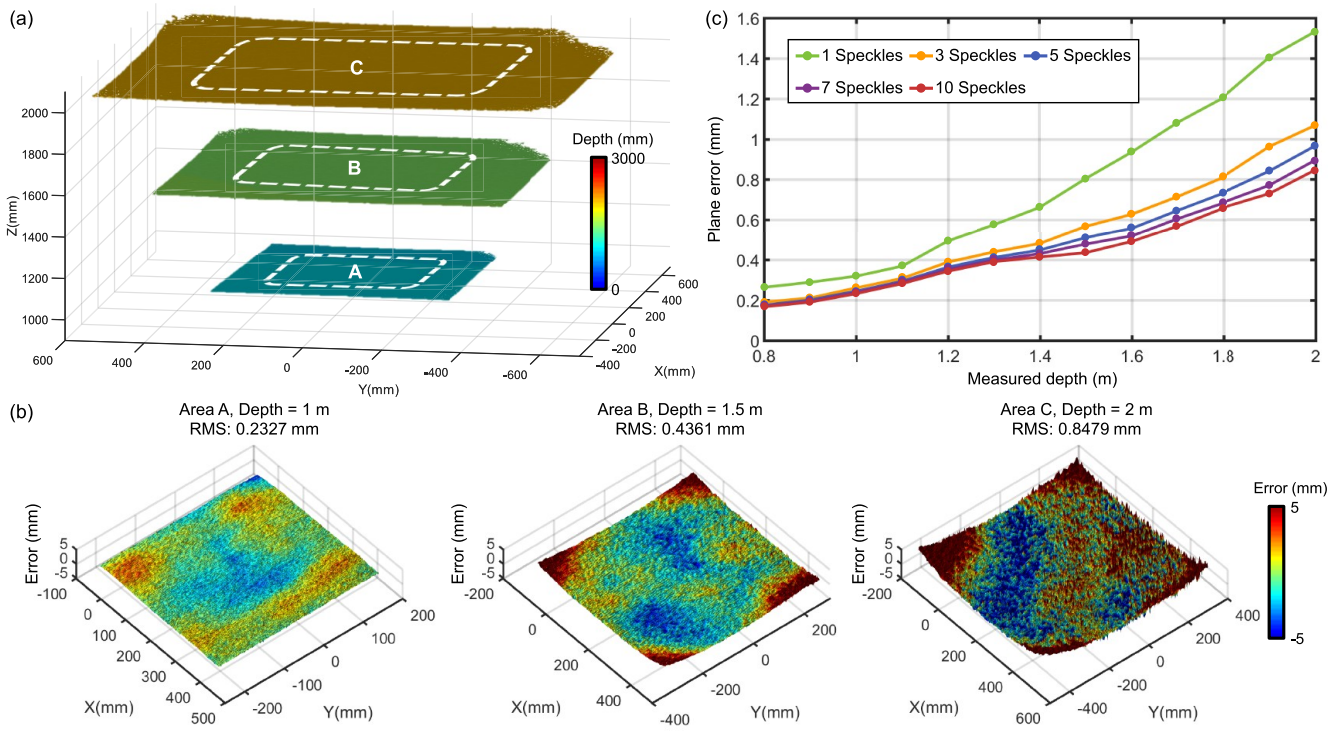


Figure 5. Precision analysis of the planar targets at different measurement distances. (a) 3D reconstruction results using ST-SPP at different distances (1, 1.5, and 2 m). (b) The error distributions of the planar targets. (c) Precision analysis results of the planar targets using different speckle patterns at distances ranging from 0.8 to 2 m.

achieved for measuring complex surfaces by projecting 10 speckle patterns, which will be applied to high-quality 3D imaging under harsh measurement conditions in industry scenarios.

EXPERIMENTS

In order to systematically verify the actual performance of the proposed ST-SPP, some experiments were carried out, including large-scale 3D reconstruction of the David model, quantitative accuracy analysis of long-range 3D imaging, dynamic scene measurement, and applications in industrial scenarios. Considering the geometric parameters of the developed 3D imaging sensor using ST-SPP, the disparity range was suitably set as −400 pixels to 119 pixels to measure scenes within a depth range from 0.8 to 2 m. In census-based spatial-temporal matching, census transform with a 5 × 5 × 10 spatial-temporal window was adopted to obtain initial 3D reconstruction results with high spatial resolution but low precision. For disparity optimization using Stereo-DIC-based subpixel spatial-temporal matching, the spatial-temporal window within the second-order shape function was enlarged to 9 × 9 × 10 after an exhaustive empirical search, further improving the matching accuracy while maintaining high spatial resolution. In addition, benefiting from the local memory mechanism and multiple operating synchronizations on the OpenCL environment, the optimized ST-SPP algorithm was parallelly computed based on GPU acceleration, which took about 27.15 ms on an NVIDIA GeForce RTX4090 graphics card, enabling high-resolution, wide-field-of-view, and real-time 3D imaging.

Large-Scale 3D Reconstruction of the David Model. First, a David plaster with a size of 40 cm (length) × 25 cm (width) × 70 cm (height) was measured to reveal the 3D imaging process of ST-SPP, and spatial-temporal speckle images captured from the left camera are shown in Figure 3a. Figure 3b and c represent a dense disparity map and ZNCC-based similarity output by ST-SPP, which can be calculated according to the equation of $C_{ZNCC} = 1-0.5 \times C_{ZNSSD}$.[63] It can be found from the similarity map that the matched points with ZNCC greater than 0.8 account for 86.24% of all valid points, demonstrating the high correctness of spatial-temporal speckle matching. To verify the availability of Stereo-DIC-based subpixel matching for complex surfaces, different subpixel disparity optimization methods were implemented for comparison. According to eq 6, the 3D reconstruction result obtained using subpixel disparity estimation based on parabola fitting is shown in Figure 3d, which can improve the disparity accuracy straightforwardly, but it yields coarse results with low quality in terms of accuracy and resolution, precluding the recovery of fine surfaces including eyes, hair, both cheeks, etc. On the contrary, the proposed ST-SPP introduces Stereo-DIC into spatial-temporal subpixel matching to overcome the matching errors caused by the perspective differences between stereo cameras, enabling high-precision 3D modeling of complex shapes as shown in Figure 3f.

Additionally, to quantitatively analyze the measurement errors, the obtained 3D point clouds were compared with the ground truth using Iterative Closest Point (ICP).[67] The high-quality ground truth was generated using a blue-ray industrial-grade 3D scanner instrument (OKIO 3M) based on FPP, which combines multiple sets of multi-step phase-shifting fringe patterns and global point cloud registration based on active targets to achieve high-precision 360° overall 3D reconstruction with an accuracy of 0.01 mm. We collected the standard 3D point cloud data of the Lu Xun model and a car door by the same means in the next experiments. Figure 3e and g show the corresponding error distributions of the measured 3D results,

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

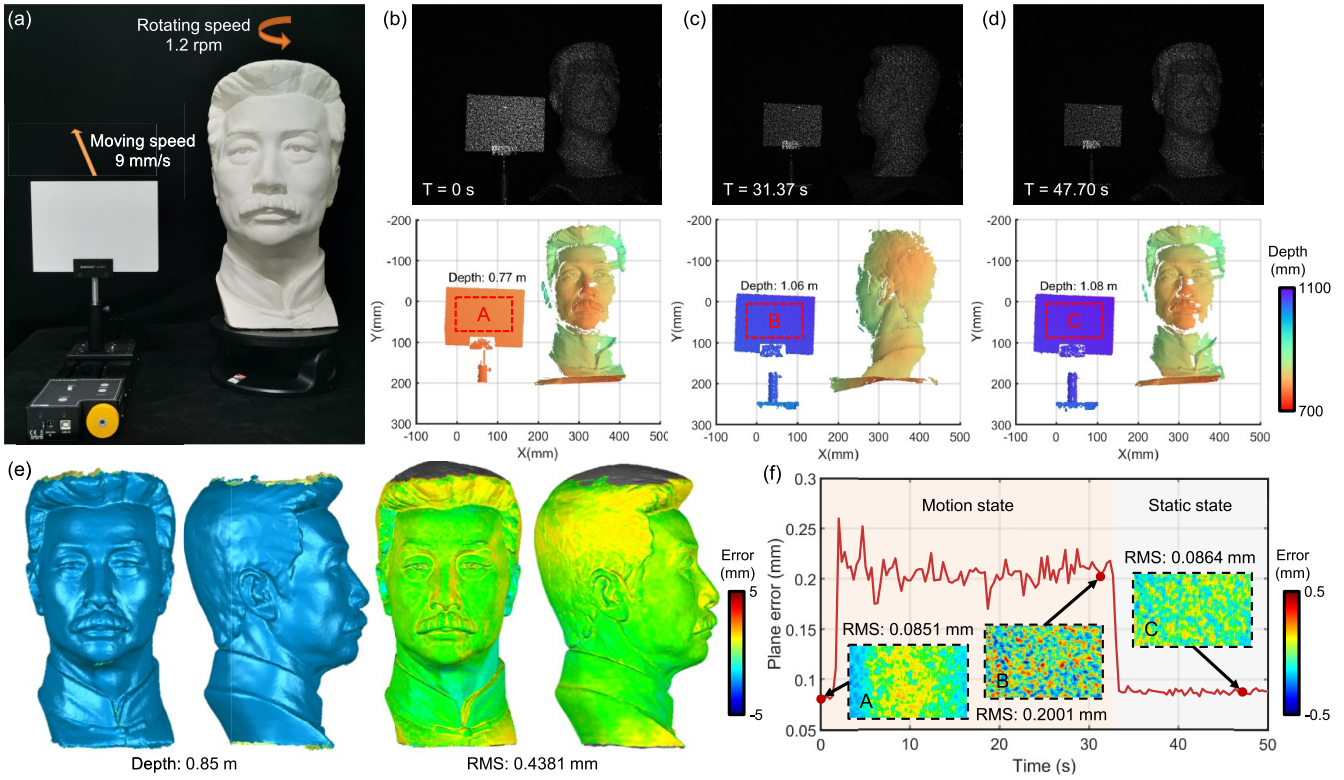


Figure 6. Dynamic 3D measurement results including a rotated Lu Xun model and a ceramic plane moving along the *z* axis. (a) The Lu Xun model and a ceramic plane. (b–d) Spatial-temporal speckle images and the corresponding 3D reconstruction results at different time points, (e) 360° 3D modeling of the Lu Xun model and the corresponding measurement errors after point cloud fusion, and (f) temporal precision analysis results of the moving plane over a 50 s period.

where major measured errors are less than 3 mm with an RMS of 0.4582 and 0.3324 mm. The magnified maps of 3D reconstruction results and measured errors in Figure 3h indicate again that the proposed ST-SPP is able to reconstruct high-quality 3D information for local fine details of objects with complex surfaces.

Next, the David model was reconstructed using different speckle patterns to verify the conclusion of the simulation experiment in Figure 4. Similarly, projecting a single speckle pattern cannot accurately recover 3D profiles of the David model with steep and ridged surfaces, and thus the obtained rough 3D measurement results with 181 638 POIs only account for 45.41% of all valid points in Figure 4a and f. Once the projected speckle patterns are increased from one frame to three images, the completeness and accuracy of 3D measurement results are greatly enhanced for removing a large number of mismatches while reducing the measurement errors to 0.4361 mm in Figure 4g. The proposed ST-SPP using five or seven speckle patterns can already provide acceptable and accurate 3D measurement results with an RMS of 0.3457 or 0.3382 mm, but its robustness under some harsh conditions needs to be boosted for the highlighted regions as shown in Figure 4h,i. Finally, experimental results similar to simulation experiments show that projecting 10 speckle images is an optimal measurement strategy with excellent performance to achieve high-precision and high-resolution 3D imaging for large-scale objects with complex profiles.

Quantitative Accuracy Analysis of Long-Range 3D Imaging. Benefiting from the large depth of field of a well-designed camera lens and the long probe distance of the speckle projection module based on high-power VCSEL, the measurement range of the proposed 3D imaging sensor is from 0.8 to 2 m, and the corresponding valid 3D imaging region is from 0.6 × 0.5 at 0.8 m to 1 × 1.2 at 2 m. To quantitatively evaluate the accuracy of long-range 3D imaging using ST-SPP, an experiment was conducted on the test scene consisting of planar targets at distances ranging from 0.8 to 2 m. Figure 5a shows the color-coded 3D reconstruction results by ST-SPP at different distances (1, 1.5, and 2 m), and plane fitting was performed on 50% of the valid areas to obtain the corresponding error distributions, where major measured errors are less than 5 mm with an RMS of 0.2376, 0.4365, and 0.8482 mm as shown in Figure 5b. Quantitative analysis results of planar targets at different distances show that ST-SPP can successfully achieve accurate and long-range 3D reconstruction with a relative accuracy higher than 0.05%. In Figure 5c, we additionally provide the measurement results using ST-SPP with different speckle patterns. As the number of speckle patterns used decreases, the measurement errors gradually deteriorate, but the measurement accuracy remains at a high level even when only using three speckle patterns, demonstrating long-range 3D imaging with the high performance of ST-SPP both in terms of measurement accuracy and efficiency.

Dynamic Scene Measurement. Next, our developed 3D sensor using ST-SPP was applied to record a dynamic scene to further validate its capability of real-time 3D shape measurement, as shown in Figure 6. The tested scene consists of a Lu Xun model with a rotating speed of 1.2 rpm and a ceramic plane moving along the *z* axis at 9 mm/s, as shown in Figure 6a. During this experiment, the exposure time of the stereo IR camera and the switching period of 10 speckle projection modules were set at 32 and 1 ms to continuously capture spatial-temporal speckle

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

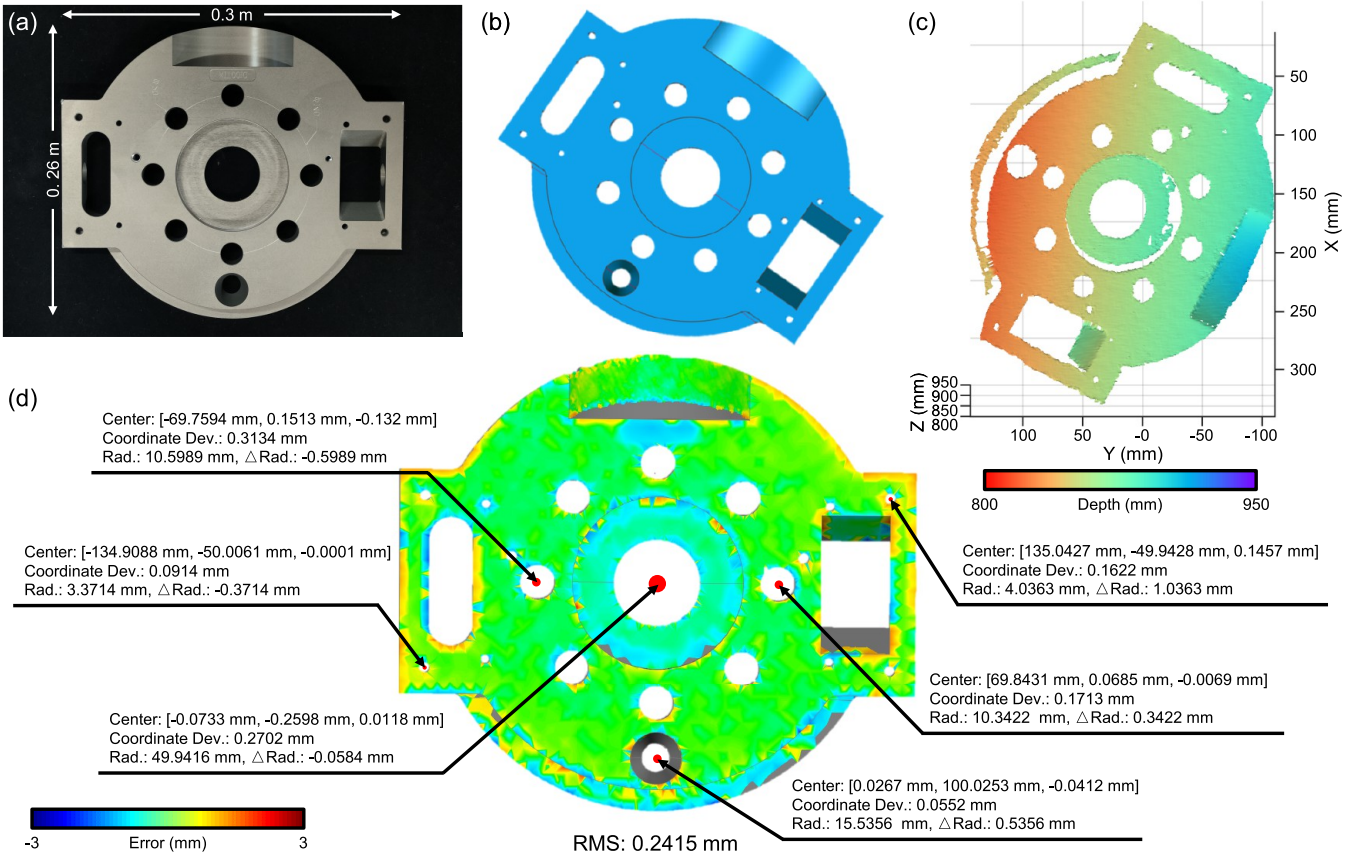


Figure 7. 3D reconstruction results of the standard industrial part. (a−d) The standard industrial part, the CAD model, 3D reconstruction results, and the corresponding measurement errors using ST-SPP.

images at a speed of 30 Hz. In order to maximize the utilization of spatial-temporal speckle images and improve the 3D measurement efficiency of ST-SPP, we adopted the sliding temporal window technique to establish a stack of speckle image sequences to dynamically process the new 10 speckle images, thereby updating a new 3D reconstruction result at 30 frames per second.

Figure 6b−d show representative spatial-temporal speckle images and corresponding color-coded 3D reconstruction results at different time points, also see Video S1. For the rotated Lu Xun model at 0.85 m, the reconstructed 3D point clouds demonstrate that our method can achieve robust 3D shape measurement of moving objects with ridged and complex surfaces. Through multiple measurements of the Lu Xun model from different views, the corresponding 3D data were further processed to merge multiple scans together for 360° 3D modeling by point cloud fusion based on ICP. The obtained 360° overall 3D reconstruction result of the Lu Xun model was compared with the ground truth to quantitatively analyze the measurement errors as shown in Figure 6e, where major measured errors are less than 5 mm with an RMS of 0.4381 mm, verifying that the proposed ST-SPP can be applied for high-quality and efficient 3D modeling of complex and large-sized objects.

In addition, we further performed temporal precision analysis of the moving plane by collecting long-term data over a 50 s period using ST-SPP. During the whole dynamic measurement, the ceramic plane moved from 0.77 to 1.08 m along the z axis at 9 mm/s, which reached the boundary of the electric slide at about 33 s and stopped until the end of the measurement. For the highlighted regions in Figure 6b−d, the error distributions of the moving plane are calculated by plane fitting at T = 0, 31.37, and 47.70 s, where major measured errors are less than 0.5 mm with an RMS of 0.0851 mm, 0.2001 m, and 0.0864 mm as shown in Figure 6f. Quantitative analysis results of the moving plane show that the measured results obtained by ST-SPP exhibit high 3D reconstruction accuracy with a low temporal standard deviation (STD) of 0.1991 ± 0.0238 mm in a motion state (from 1 to 33 s) and 0.0886 ± 0.0020 mm under static conditions (from 33 to 50 s). Temporal precision analysis results show that the measurement errors caused by motion artifacts lead to a small decline in the plane measurement accuracy but remain at a high level. These experimental results suggest that ST-SPP is a promising tool for high-resolution, wide-field-of-view, and real-time 3D measurements for dynamic scenes with complex shapes.

Applications in Industrial Scenarios. Last of all, to verify the applicability of ST-SPP in industrial scenarios, we tested a standard industrial part manufactured using the designed CAD model with a size of 0.3 × 0.26 m (length) × 0.26 m (width). The corresponding 3D reconstruction results using ST-SPP in Figure 7a−c confirm that the measurement errors are smaller in smooth planar regions but larger in sharp edges, and the RMS is 0.2415 mm as shown in Figure 7d. Further, we adopted the CAD model as the benchmark to calculate the position errors of some key points in the measured 3D point clouds. The center coordinates and radii of seven selected circles were measured through circle fitting, and the obtained coordinate deviations range from 0.0552 to 0.3134 mm compared with the reference values. These circles are custom-made with radii varying from 3, 10, 15, and 50 mm. It can be found that the smaller the radius of

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

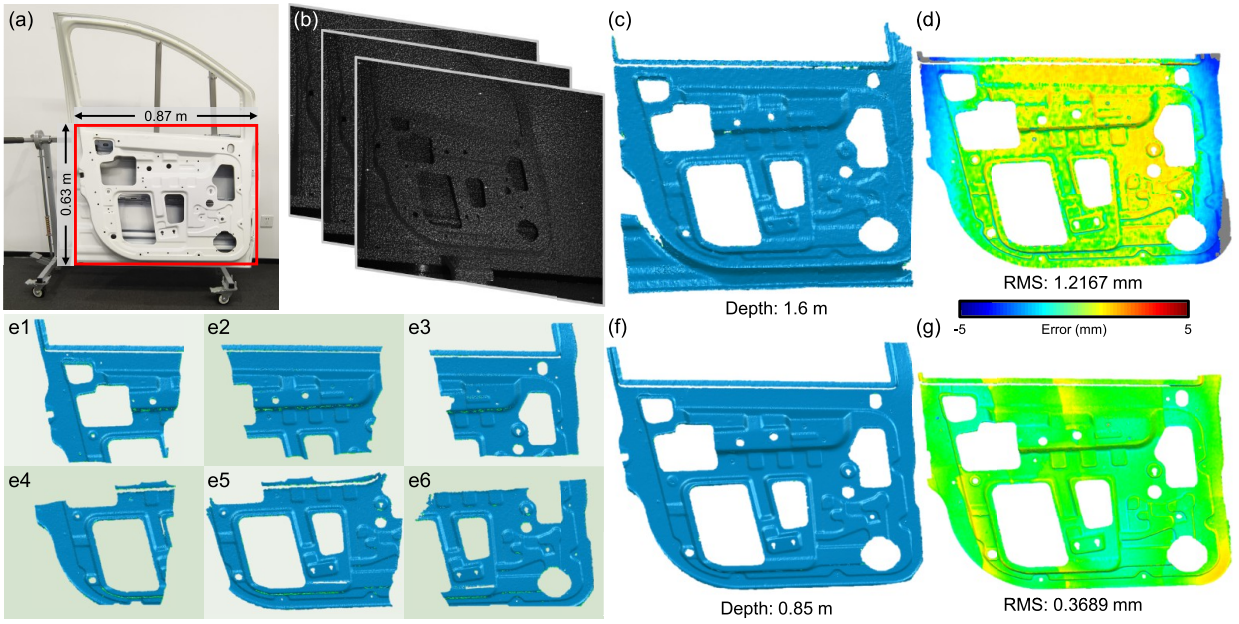


Figure 8. 3D reconstruction results of the car door. (a,b) The car door and spatial-temporal speckle images. (c,d) The single 3D reconstruction and the corresponding measurement errors at 1.6 m. (e1−e6) Multiple local 3D reconstruction results at 0.85 m. (f,g) 3D reconstruction results and the corresponding measurement errors after point cloud fusion.

the circle, the poorer the measured radius deviation. Due to the limited image resolution of the camera, the X and Y resolution of our 3D imaging sensor is about 0.43 mm at 0.8 m. The proposed ST-SPP cannot accurately measure the circle with a radius of 3 mm but enables precision positioning with a radius deviation of better than 0.6 mm for circles with other sizes. These quantitative analysis results demonstrate that our method can be applied for high-quality and efficient 3D modeling of complex structure parts as shown in Figure 7c,d.

Next, to further demonstrate the advantages of ST-SPP, we applied the proposed long-range 3D imaging sensor to the reconstruction of a car door with a size of 0.87 m (length) × 0.63 m (width), as shown in Figure 8a,b. First, a global and coarse 3D point cloud of a large-scale car door was obtained by a single measurement at 1.6 m in Figure 8c, and the corresponding error distribution was calculated by comparing with the ground truth, where major measured errors are less than 5 mm with an RMS of 1.2167 mm, as shown in Figure 8d. Multiple close-range measurements were then performed at 0.85 m to recover local fine details of the car door, and the corresponding multiple sets of 3D reconstruction results were guided by the global 3D data to achieve robust and efficient automatic point cloud fusion, as shown in Figure 8e,f. Figure 8g shows that combining ST-SPP with multiple global and local measurements enables high-precision and fast 3D shape measurement with high completeness for large-sized objects, reducing the RMS from 1.2167 to 0.3698 mm. These results suggest that the proposed 3D imaging sensor based on ST-SPP can be applied for long-range 3D measurement and reverse modeling with high quality for large-scale objects with complex shapes, exhibiting the advantages of its application in industrial inspection, workpiece assembly, and other fields.

CONCLUSIONS AND DISCUSSIONS

In summary, we have demonstrated a miniaturized and low-cost 3D imaging sensor based on spatial-temporal speckle projection profilometry (ST-SPP) with a VCSEL projector array that can achieve high-resolution, wide-field-of-view, and real-time 3D measurement. By integrating a set of VCSEL-based speckle projectors, ST-SPP performs spatial-temporal encoding with global uniqueness for the depth information on the measured scenes, overcoming the degradation of reconstruction quality in single-frame SPP methods. Utilizing the proposed coarse-to-fine spatial-temporal matching strategy using DIC, ST-SPP strengthens the ability of speckle matching for complex surfaces to improve the accuracy of subpixel disparity estimation. The effectiveness of ST-SPP has been verified by several experiments for measuring various types of large-size samples. First, a virtual environment was created using Blender to simulate the performance of the developed 3D sensor, which proved that projecting 10 speckle images is an optimal measurement strategy to achieve more robust 3D imaging with the highest accuracy for complex surfaces. The measurement of large-scale David confirmed that ST-SPP can achieve high-quality 3D reconstruction with an accuracy improvement from 0.4582 to 0.3533 mm using Stereo-DIC-based subpixel matching compared with traditional subpixel fitting. Precision analysis results verified the high accuracy and efficiency of ST-SPP for measuring long-range planar targets, and the relative accuracy is higher than 0.05% at distances ranging from 0.8 to 2 m even when only using three speckle patterns. For dynamic scene measurement, 360° reconstruction of the Lu Xun model and precision analysis of a moving ceramic plane confirmed the advantages of ST-SPP for real-time and high-precision 3D imaging with a low temporal standard deviation of 0.1991 ± 0.0238 mm in motion state and 0.0886 ± 0.0020 mm under static conditions. Finally, 3D reconstruction of an industrial part and a car door revealed the capability of ST-SPP for accurate and efficient large-scale reverse 3D modeling and precision positioning, exhibiting the advantages of its industrial application in quality inspection and workpiece assembly.

We have discussed the 3D imaging performance of the proposed ST-SPP based on spatial-temporal speckle projection by integrating a set of VCSEL-based miniaturized speckle

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511

projectors. In the future, we will enhance the measurement accuracy and spatial resolution of ST-SPP to successfully run on an embedded low-power platform for mobile applications in optical metrology, further pushing the limits of ST-SPP in speed, accuracy, resolution, measurement range, and portability. Recently, compared with traditional stereo matching methods, many deep learning methods for stereo vision have been proposed, which have achieved excellent performance of stereo matching.[68−72] How to build a lightweight spatial-temporal speckle matching network based on Stereo-DIC to strengthen the measurement accuracy and resolution at the cost of a low computational overhead while using fewer speckle patterns is an interesting direction for further research. The related applications of ST-SPP will be explored in state-of-the-art manufacturing processes, precision positioning, and quality assessment, realizing the leap from 3D visual guidance to 3D visual measurement and then to 3D visual inspection.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphotonics.3c01341.

Supplemental document: Additional information on VCSEL-based speckle projection and the relationship between the slanted window model in PatchMatch and the shape function in Stereo-DIC (PDF)

Video S1: Dynamic 3D measurement results including a rotated Lu Xun model and a ceramic plane moving along the z axis (MP4)

■ AUTHOR INFORMATION

Corresponding Authors

Shijie Feng − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*; Email: shijiefeng@njust.edu.cn

Chao Zuo − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*; orcid.org/0000-0002-1461-0032; Email: zuochao@njust.edu.cn

Authors

Wei Yin − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

Hang Zhao − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

Yifan Ji − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

ZhongDa Deng − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

Ziheng Jin − *Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China; Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China; Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

Xiaolei Zhang − *Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China*

Huai Wang − *Suzhou Abham Intelligent Technology Co., Ltd., Suzhou, Jiangsu Province 215000, China*

Qian Chen − *Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphotonics.3c01341

Author Contributions

⊥These authors contributed equally to this work.

Funding

This work was supported by National Key Research and Development Program of China (2022YFB2804603, 2022YFB2804604), National Natural Science Foundation of China (62205147, 62075096, U21B2033), China Postdoctoral Science Foundation (2023T160318, 2022M711630, 2022M721619), Jiangsu Funding Program for Excellent Postdoctoral Talent (2022ZB254), the Leading Technology of Jiangsu Basic Research Plan (BK20192003), the "333 Engineering" Research Project of Jiangsu Province (BRA2016407), the Jiangsu Provincial "One Belt and One Road" innovation cooperation project (BZ2020007), Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense (JSGP202105), Fundamental Research Funds for the Central

Universities (30922010405, 30921011208, 30920032101, 30919011222), National Major Scientific Instrument Development Project (62227818), and Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_0437).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thanks the editor and reviewers for their help on this paper.

■ REFERENCES

(1) Gåsvik, K. J. *Optical Metrology*; John Wiley & Sons, 2003.
(2) Zuo, C.; Qian, J.; Feng, S.; Yin, W.; Li, Y.; Fan, P.; Han, J.; Qian, K.; Chen, Q. Deep learning in optical metrology: a review. *Light: Sci. Appl.* **2022**, *11*, 1−54.
(3) Wyant, J. C.; Creath, K. Recent advances in interferometric optical testing. *Laser Focus* **1985**, *21*, 118−132.
(4) Kulkarni, R.; Rastogi, P. Optical measurement techniques-A push for digitization. *Optics and Lasers in Engineering* **2016**, *87*, 1−17.
(5) Wu, Z.; Kang, I.; Yao, Y.; Jiang, Y.; Deng, J.; Klug, J.; Vogt, S.; Barbastathis, G. Three-dimensional nanoscale reduced-angle ptycho-tomographic imaging with deep learning (RAPID). *eLight* **2023**, *3* (1), 1−12.
(6) Liu, C.; Jiang, Z.; Wang, X.; Zheng, Y.; Zheng, Y.-W.; Wang, Q.-H. Continuous optical zoom microscope with extended depth of field and 3D reconstruction. *PhotoniX* **2022**, *3* (1), 20.
(7) Huang, Z.; Memmolo, P.; Ferraro, P.; Cao, L. Dual-plane coupled phase retrieval for non-prior holographic imaging. *PhotoniX* **2022**, *3* (1), 1−16.
(8) Ganapathi, V.; Plagemann, C.; Koller, D.; Thrun, S. Real time motion capture using a single time-of-flight camera. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; IEEE, 2010; pp 755−762.
(9) Kolb, A.; Barth, E.; Koch, R.; Larsen, R. Time-of-flight cameras in computer graphics. *Computer Graphics Forum* **2010**, *29*, 141−159.
(10) Hansard, M.; Lee, S.; Choi, O.; Horaud, R. P. *Time-of-flight cameras: principles, methods and applications*; Springer Science & Business Media, 2012.
(11) Sun, J.; Zheng, N.-N.; Shum, H.-Y. Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2003**, *25*, 787−800.
(12) Lazaros, N.; Sirakoulis, G. C.; Gasteratos, A. Review of stereo vision algorithms: from software to hardware. *International Journal of Optomechatronics* **2008**, *2*, 435−462.
(13) Gorthi, S. S.; Rastogi, P. Fringe projection techniques: whither we are? *Opt. Laser Eng.* **2010**, *48*, 133−140.
(14) Feng, S.; Zhang, L.; Zuo, C.; Tao, T.; Chen, Q.; Gu, G. High dynamic range 3D measurements with fringe projection profilometry: a review. *Mea. Sci. Technol.* **2018**, *29*, No. 122001.
(15) Zhang, S. Absolute phase retrieval methods for digital fringe projection profilometry: A review. *Opt. Laser Eng.* **2018**, *107*, 28−37.
(16) Yin, W.; Feng, S.; Tao, T.; Huang, L.; Zhang, S.; Chen, Q.; Zuo, C. Calibration method for panoramic 3D shape measurement with plane mirrors. *Opt. Express* **2019**, *27*, 36538−36550.
(17) Guo, W.; Wu, Z.; Li, Y.; Liu, Y.; Zhang, Q. Real-time 3D shape measurement with dual-frequency composite grating and motion-induced error reduction. *Opt. Express* **2020**, *28*, 26882−26897.
(18) Wu, Z.; Guo, W.; Zhang, Q.; Wang, H.; Li, X.; Chen, Z. Time-overlapping structured-light projection: high performance on 3D shape measurement for complex dynamic scenes. *Opt. Express* **2022**, *30*, 22467−22486.
(19) Kim, G.; Kim, Y.; Yun, J.; Moon, S.-W.; Kim, S.; Kim, J.; Park, J.; Badloe, T.; Kim, I.; Rho, J. Metasurface-driven full-space structured light for three-dimensional imaging. *Nature Communications* **2022**, *13* (1), 5920.
(20) Jing, X.; Zhao, R.; Li, X.; Jiang, Q.; Li, C.; Geng, G.; Li, J.; Wang, Y.; Huang, L. Single-shot 3D imaging with point cloud projection based on metadevice. *Nature Communications* **2022**, *13* (1), 7842.
(21) Wolcott, R. W.; Eustice, R. M. Visual localization within lidar maps for automated urban driving. *2014 IEEE/RSJ. International Conference on Intelligent Robots and Systems*; IEEE, 2014; pp 176−183.
(22) Hirschmüller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2008**, *30*, 328−341.
(23) Hirschmüller, H.; Scharstein, D. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2009**, *31*, 1582−1599.
(24) Geiger, A.; Roser, M.; Urtasun, R. Efficient large-scale stereo matching. *Asian Conference on Computer Vision* **2011**, *6492*, 25−38.
(25) Mei, X.; Sun, X.; Zhou, M.; Jiao, S.; Wang, H.; Zhang, X. On building an accurate stereo matching system on graphics hardware. *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*; IEEE, 2011; pp 467−474.
(26) Bleyer, M.; Rhemann, C.; Rother, C. PatchMatch Stereo-Stereo Matching with Slanted Support Windows. *BMVC* **2011**, 1−11.
(27) Zhang, Z. Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques. *Opt. Laser Eng.* **2012**, *50*, 1097−1106.
(28) Zuo, C.; Tao, T.; Feng, S.; Huang, L.; Asundi, A.; Chen, Q. Micro Fourier transform profilometry (μFTP): 3D shape measurement at 10,000 frames per second. *Opt. Laser Eng.* **2018**, *102*, 70−91.
(29) Zhang, S. High-speed 3D shape measurement with structured light methods: A review. *Opt. Laser Eng.* **2018**, *106*, 119−131.
(30) Heist, S.; Dietrich, P.; Landmann, M.; Kühmstedt, P.; Notni, G.; Tünnermann, A. GOBO projection for 3D measurements at highest frame rates: a performance analysis. *Light: Sci. Appl.* **2018**, *7*, 1−13.
(31) Yin, W.; Zuo, C.; Feng, S.; Tao, T.; Hu, Y.; Huang, L.; Ma, J.; Chen, Q. High-speed three-dimensional shape measurement using geometry-constraint-based number-theoretical phase unwrapping. *Opt. Laser Eng.* **2019**, *115*, 21−31.
(32) Schaffer, M.; Grosse, M.; Kowarschik, R. High-speed pattern projection for three-dimensional shape measurement using laser speckles. *Appl. Opt.* **2010**, *49*, 3622−3629.
(33) Schaffer, M.; Grosse, M.; Harendt, B.; Kowarschik, R. High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection. *Opt. Lett.* **2011**, *36*, 3097−3099.
(34) Zhou, P.; Zhu, J.; Jing, H. Optical 3-D surface reconstruction with color binary speckle pattern encoding. *Opt. Express* **2018**, *26*, 3452−3465.
(35) Yin, W.; Hu, Y.; Feng, S.; Huang, L.; Kemao, Q.; Chen, Q.; Zuo, C. Single-shot 3D shape measurement using an end-to-end stereo matching network for speckle projection profilometry. *Opt. Express* **2021**, *29*, 13388−13407.
(36) Su, X.; Chen, W. Fourier transform profilometry: a review. *Opt. Laser Eng.* **2001**, *35*, 263−284.
(37) Feng, S.; Chen, Q.; Gu, G.; Tao, T.; Zhang, L.; Hu, Y.; Yin, W.; Zuo, C. Fringe pattern analysis using deep learning. *Adv. Photonics* **2019**, *1*, No. 025001.
(38) Yin, W.; Zhong, J.; Feng, S.; Tao, T.; Han, J.; Huang, L.; Chen, Q.; Zuo, C. Composite deep learning framework for absolute 3D shape measurement based on single fringe phase retrieval and speckle correlation. *Journal of Physics: Photonics* **2020**, *2*, No. 045009.
(39) Zuo, C.; Feng, S.; Huang, L.; Tao, T.; Yin, W.; Chen, Q. Phase shifting algorithms for fringe projection profilometry: A review. *Opt. Laser Eng.* **2018**, *109*, 23−59.
(40) Feng, S.; Zuo, C.; Zhang, L.; Yin, W.; Chen, Q. Generalized framework for non-sinusoidal fringe analysis using deep learning. *Photonics Research* **2021**, *9*, 1084−1098.
(41) Yin, W.; Che, Y.; Li, X.; Li, M.; Hu, Y.; Feng, S.; Lam, E. Y.; Chen, Q.; Zuo, C. Physics-informed deep learning for fringe pattern analysis. *OEA* **2024**, *7*, 230034.
(42) Su, X.; Chen, W. Reliability-guided phase unwrapping algorithm: a review. *Opt. Laser Eng.* **2004**, *42*, 245−261.

(43) Zhong, K.; Li, Z.; Shi, Y.; Wang, C.; Lei, Y. Fast phase measurement profilometry for arbitrary shape objects without phase unwrapping. *Opt. Laser Eng.* **2013**, *51*, 1213−1222.
(44) Zuo, C.; Huang, L.; Zhang, M.; Chen, Q.; Asundi, A. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. *Opt. Laser Eng.* **2016**, *85*, 84−103.
(45) Wu, Z.; Guo, W.; Zhang, Q. High-speed three-dimensional shape measurement based on shifting Gray-code light. *Opt. Express* **2019**, *27*, 22631−22644.
(46) Yin, W.; Feng, S.; Tao, T.; Huang, L.; Trusiak, M.; Chen, Q.; Zuo, C. High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system. *Opt. Express* **2019**, *27*, 2411−2431.
(47) Yin, W.; Chen, Q.; Feng, S.; Tao, T.; Huang, L.; Trusiak, M.; Asundi, A.; Zuo, C. Temporal phase unwrapping using deep learning. *Sci. Reports* **2019**, *9*, 1−12.
(48) Su, X.; Zhang, Q. Dynamic 3-D shape measurement method: a review. *Opt. Laser Eng.* **2010**, *48*, 191−204.
(49) Feng, S.; Zuo, C.; Tao, T.; Hu, Y.; Zhang, M.; Chen, Q.; Gu, G. Robust dynamic 3-d measurements with motion-compensated phase-shifting profilometry. *Opt. Laser Eng.* **2018**, *103*, 127−138.
(50) Salvi, J.; Pages, J.; Batlle, J. Pattern codification strategies in structured light systems. *Pattern Recognition* **2004**, *37*, 827−849.
(51) Zhou, P.; Zhu, J.; Xiong, W.; Zhang, J. 3D face imaging with the spatial-temporal correlation method using a rotary speckle projector. *Appl. Opt.* **2021**, *60*, 5925−5935.
(52) Xiong, W.; Yang, H.; Zhou, P.; Fu, K.; Zhu, J. Spatiotemporal correlation-based accurate 3D face imaging using speckle projection and real-time improvement. *Applied Sciences* **2021**, *11*, 8588.
(53) Zhou, P.; Zhu, J.; You, Z. 3-D face registration solution with speckle encoding based spatial-temporal logical correlation algorithm. *Opt. Express* **2019**, *27*, 21004−21019.
(54) Min, D.; Lu, J.; Do, M. N. A revisit to cost aggregation in stereo matching: How far can we reduce its computational redundancy? *2011 International Conference on Computer Vision*; IEEE, 2011; pp 1567−1574.
(55) Scharstein, D.; Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* **2002**, *47*, 7−42.
(56) Fu, K.; Xie, Y.; Jing, H.; Zhu, J. Fast spatial-temporal stereo matching for 3D face reconstruction under speckle pattern projection. *Image and Vision Computing* **2019**, *85*, 36−45.
(57) Tang, Q.; Liu, C.; Cai, Z.; Zhao, H.; Liu, X.; Peng, X. An improved spatiotemporal correlation method for high-accuracy random speckle 3D reconstruction. *Opt. Laser Eng.* **2018**, *110*, 54−62.
(58) Heist, S.; Lutzke, P.; Schmidt, I.; Dietrich, P.; Kühmstedt, P.; Tünnermann, A.; Notni, G. High-speed three-dimensional shape measurement using GOBO projection. *Optics and Lasers in Engineering* **2016**, *87*, 90−96.
(59) Dietrich, P.; Heist, S.; Landmann, M.; Kühmstedt, P.; Notni, G. BICOS−an algorithm for fast real-time correspondence search for statistical pattern projection-based active stereo sensors. *Applied Sciences* **2019**, *9*, 3330.
(60) Gu, F.; Song, Z.; Zhao, Z. Single-Shot Structured Light Sensor for 3D Dense and Dynamic Reconstruction. *Sensors* **2020**, *20*, 1094.
(61) Yin, W.; Cao, L.; Zhao, H.; Hu, Y.; Feng, S.; Zhang, X.; Shen, D.; Wang, H.; Chen, Q.; Zuo, C. Real-time and accurate monocular 3D sensor using the reference plane calibration and an optimized SGM based on opencl acceleration. *Optics and Lasers in Engineering* **2023**, *165*, No. 107536.
(62) Sutton, M.; Hild, F. Recent advances and perspectives in digital image correlation. *Experimental Mechanics* **2015**, *55*, 1−8.
(63) Pan, B. Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals. *Measurement Science and Technology* **2018**, *29*, No. 082001.
(64) Gao, Y.; Cheng, T.; Su, Y.; Xu, X.; Zhang, Y.; Zhang, Q. High-efficiency and high-accuracy digital image correlation for three-dimensional measurement. *Optics and Lasers in Engineering* **2015**, *65*, 73−80.
(65) Pan, B.; Li, K.; Tong, W. Fast, robust and accurate digital image correlation calculation without redundant computations. *Experimental Mechanics* **2013**, *53*, 1277−1289.
(66) Yin, W.; Ji, Y.; Chen, J.; Li, R.; Feng, S.; Chen, Q.; Pan, B.; Jiang, Z.; Zuo, C. Initializing and accelerating Stereo-DIC computation using semi-global matching with geometric constraints. *Optics and Lasers in Engineering* **2024**, *172*, No. 107879.
(67) Mohammadzade, H.; Hatzinakos, D. Iterative closest normal point for 3D face recognition. *IEEE transactions on pattern analysis and machine intelligence* **2013**, *35*, 381−397.
(68) Zbontar, J.; LeCun, Y. Computing the stereo matching cost with a convolutional neural network. *2015 IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2015; pp 1592−1599.
(69) Luo, W.; Schwing, A. G.; Urtasun, R. Efficient deep learning for stereo matching. *2016 IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2016; pp 5695−5703.
(70) Kendall, A.; Martirosyan, H.; Dasgupta, S.; Henry, P.; Kennedy, R.; Bachrach, A.; Bry, A. End-to-end learning of geometry and context for deep stereo regression. *2017 IEEE Conference on International Conference on Computer Vision*; IEEE, 2017; 66−75.
(71) Khamis, S.; Fanello, S.; Rhemann, C.; Kowdle, A.; Valentin, J.; Izadi, S. Stereonet: Guided hierarchical refinement for real-time edge-aware depth prediction. *2018 IEEE Conference on European Conference on Computer Vision (ECCV)*; Springer: Cham, 2018; vol *11219*, pp 596−613..
(72) Chang, J.-R.; Chen, Y.-S. Pyramid stereo matching network. *2018 IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2018; pp 5410−5418.

https://doi.org/10.1021/acsphotonics.3c01341
ACS Photonics 2024, 11, 498−511