Optics and Lasers in Engineering 172 (2024) 107879

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Initializing and accelerating Stereo-DIC computation using semi-global matching with geometric constraints

Wei Yin [a,b,c,1], Yifan Ji [a,b,c,1], Juntong Chen [a,b,c], Rui Li [d], Shijie Feng [a,b,c], Qian Chen [c], Bing Pan [e,*], Zhenyu Jiang [d,*], Chao Zuo [a,b,c,*]

[a] Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China
[b] Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China
[c] Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing, Jiangsu Province 210094, China
[d] State Key Laboratory of Subtropical Building Science, School of Civil Engineering and Transportation, South China University of Technology, Guangzhou, Guangdong Province 510640, China
[e] National Key Laboratory of Strength and Structural Integrity, School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China

ARTICLE INFO

Keywords:
Digital Image Correlation
3D imaging
Stereo matching
3D deformation measurement

ABSTRACT

Stereo Digital Image Correlation (Stereo-DIC) has become a mainstream optical metrology technique for quantitatively analyzing full-field 3D shape, displacement, or deformation of materials and structures. Whether it is to measure 3D profile or deformation, stereo matching is essential for Stereo-DIC to reconstruct 3D point clouds from stereo images. Traditional feature-based (e.g., SIFT) methods provide initial 2D displacements for stereo matching with the aid of extracting massive features, but at the cost of expensive computational overhead. In addition, these methods preclude precise measurement of objects with steep and ridged surfaces or undergoing large rotation and/or deformation due to low feature matching accuracy of complex regions caused by perspective differences. In this paper, we propose a fast and robust stereo matching method using semi-global matching with geometric constraints (GC-SGM) for initializing and accelerating Stereo-DIC computation. For GC-SGM, an optimized semi-global matching (SGM) algorithm based on GPU acceleration is first utilized to quickly estimate dense and reliable disparity maps between the rectified stereo images. The global pixel-wise 2D correspondence between raw stereo images can be established inversely using epipolar constraints and 1D disparity information, and then converted to accurate and initial second-order deformation parameters for 2D-DIC-based sub-pixel refinement by least-squares-based surface fitting. Experimental results prove that the proposed GC-SGM enhances the matching correctness and robustness for complex objects while improving the processing speed on GPU by 3~10 times compared with SIFT-based methods, enabling high-precision and computationally efficient 3D shape and deformation measurement.

1. Introduction

Stereo Digital Image Correlation (Stereo-DIC), as a non-contact and full-field 3D shape and deformation measurement technique based on the principles of DIC and stereo vision, has been successfully applied to various areas, such as materials science, biomechanics, and aerospace engineering [1,2]. Image correlation is pivotal to Stereo-DIC and it can be divided into two categories: temporal matching and stereo matching. Temporal matching, which can be seen as subset matching based on 2D-DIC, is to track the 2D displacement of points of interest (POIs) on an image sequence continuously acquired by one camera. 2D-DIC, as a local iterative optimization technique, requires an accurate initial guess of the deformation or displacement of valid points to initialize the iterative calculation [2]. In the common instance of temporal matching, the 2D displacement of POIs is first initialized by different methods including fast-Fourier transform (FFT) [3], genetic algorithms [4], and scale-invariant feature transform (SIFT) [5,6], and further refined using 2D-DIC with classic NR algorithms [7] or inverse-compositional Gauss–Newton (IC-GN) algorithms [8,9]. FFT methods are computationally efficient and enable global sub-pixel displacement estimation by detecting

* Corresponding authors.
E-mail addresses: panb@buaa.edu.cn (B. Pan), zhenyujiang@scut.edu.cn (Z. Jiang), zuochao@njust.edu.cn (C. Zuo).
[1] They contributed equally to this work.

https://doi.org/10.1016/j.optlaseng.2023.107879
Received 8 August 2023; Received in revised form 16 September 2023; Accepted 2 October 2023
Available online 13 October 2023

peaks in complex spectra by fast Fourier transform [10]. But it suffers from the accuracy and robustness of peak-finding algorithms, resulting in the degradation of reconstruction quality for objects with complex surfaces. In contrast, SIFT is quite robust to ambient illumination and varying surface properties, enabling it to achieve higher-accuracy keypoint matching with the aid of extracting massive scale-invariant features, but its feature extraction and matching are computationally expensive, which limits the applications of the SIFT-aided DIC. Yang et al. [11] developed a SIFT-aided path-independent DIC method to estimate the initial guess of all valid points by introducing the parallel computing on the graphics processing unit (GPU) or multi-core CPU, achieving real-time processing with high resolution and accuracy for temporal matching based on 2D-DIC.

Different from temporal matching, in stereo matching, the 2D correspondence between stereo images captured by left and right cameras is established and converted into 3D shape information of the tested object after stereo calibration. However, for initial 2D displacement estimation, traditional feature-based (e.g., SIFT) methods preclude precise measurement of objects with steep and ridged surfaces or undergoing large rotation and/or deformation due to low feature matching accuracy of complex regions caused by perspective differences. Lin et al. [12] presented an epipolar constraint-aided 2D correspondence searching method to cope well with objects with relatively steep slopes. However, the running time of this method on GPU was increased by nearly 2.35 times compared with SIFT feature-aided approaches. In addition, it is worth noting that the perspective differences between stereo cameras make the first-order shape function commonly used in temporal matching no longer applicable, and a more complex second-order shape function is used for 2D-DIC-based sub-pixel refinement in stereo matching [13]. Therefore, compared with temporal matching, stereo matching is a more challenging and computationally complex part of Stereo-DIC, and its measurement accuracy and computational efficiency need to be improved urgently [14].

On the basis of 3D shape measurement using stereo matching, there are generally three matching strategies in Stereo-DIC to realize 3D deformation measurement for quantitatively analyzing the 3D displacement field of point clouds at different times or states. As shown in Figs. 1 (a)-(c), the commonality of the three strategies is that temporal matching is exploited to process the left image sequence for tracking the same subsets at the initial state and other states, and stereo matching is performed on left and right reference images for obtaining initial 3D point clouds of objects. In addition, in the first strategy, temporal matching processes the right reference image and others on the right image sequence. The second strategy is that the left reference image is matched to all frames in the right image sequence. For the last strategy, stereo matching is implemented on left and right images collected at the same time in the whole deformed process. Since stereo matching is more computationally complex than temporal matching, the first strategy that only needs to perform stereo matching once is often preferred. It can be found that stereo matching is frequently used in the third matching strategy. However, this strategy not only suffers from expensive computational costs from stereo matching, but also requires the interpolation of left images (with complex calculation and non-negligible interpolation error) at each state [2], which is not recommended in Stereo-DIC.

In this paper, the main contribution is to renovate the third matching strategy with poor performance. By introducing the adaptive subset offset scheme [15], a stereo matching method using semi-global matching with geometric constraints (GC-SGM) is proposed to markedly improve the computational efficiency and accuracy in Stereo-DIC, outputting comparable measurement results to the first matching strategy. For GC-SGM, benefiting from the shared memory mechanism and multiple stream operations on the CUDA environment, an optimized semi-global matching (SGM) algorithm based on GPU acceleration is first utilized to realize fast and global disparity estimation for the rectified stereo image. Further, the obtained disparity information is combined with epipolar constraints to inversely achieve initial 2D displacement matching between raw stereo images. For the second-order shape function involved in stereo matching, initial deformation parameters of each point are calculated using least-squares-based surface fitting, and then refined to enhance the accuracy of 2D displacement measurement using 2D-DIC based on IC-GN. Experiments proved that the proposed GC-SGM enhances the matching correctness and robustness while improving the processing speed on GPU by 3~10 times compared with SIFT-based methods, achieving fast and accurate 3D profile and deformation measurements for objects with steep and ridged surfaces or undergoing large rotation and/or deformation.

2. Principle

As shown in Fig. 1 (d), the proposed GC-SGM consists of four steps to initialize second-order deformation parameters and accelerate the computation of stereo matching in Stereo-DIC. For stereo image pairs $\{I_L, I_R\}$ captured at the same moment, epipolar rectification is first executed to align the epipolar lines of left and right images, simplifying the two-dimensional search problem to a one-dimensional matching problem. The rectified images $\{I_L^{Rect}, I_R^{Rect}\}$ are processed to obtain the 1D disparity map D^{Rect} using an optimized semi-global matching (SGM) method by parallel calculation on GPU according to our previous work [16]. Based on stereo calibration parameters and epipolar constraints, for any point $\mathbf{p} = [x_L, y_L]^\top$ on left image I_L, initial 2D displacement of the corresponding point $\mathbf{q} = [x_R, y_R]^\top$ on right image I_R can be determined using D^{Rect}, and converted to second-order deformation parameters for stereo matching using surface fitting and subset matching based on 2D-DIC.

2.1. Optimized semi-global matching (SGM) based on GPU acceleration

Stereo-DIC and stereo vision are both mainstream optical metrology [17–19] and computational imaging [20–22] techniques for measuring the 3D profile of the tested scenes. Different from subset matching commonly used in Stereo-DIC, stereo matching, as the core technology in stereo vision, is based on local matching methods, semi-global matching (SGM) methods, or global matching methods, which are implemented to build the global 1D correspondence of rectified stereo images and obtain dense disparity maps. There is generally a four-step pipeline for stereo matching, including matching cost calculation, cost aggregation, disparity computation, and disparity refinement. Due to the limited computing resource early, some GPU-accelerated local matching methods were proposed to enhance the matching accuracy and running speed using cost calculation methods based on different similarity estimation criteria or multi-scale matching windows [23–25]. However, in the well-known stereo vision dataset Middlebury, the matching accuracy of these methods is lower than that of global stereo matching methods, which can achieve accurate pixel-level stereo matching through a global energy function at the cost of higher computational complexity [26]. By inheriting the advantages of the above two methods, SGM methods have the same algorithmic complexity as local methods by accumulating one-dimensional cost aggregation results from all directions, striking a good balance in stereo matching accuracy and computational efficiency [27,28].

Here, by artificially manufacturing or projecting speckle patterns onto the surface of the measured objects to enhance their features, we proposed an optimized SGM algorithm based on GPU acceleration to overcome streak artifacts in stereo matching, achieving efficient, dense, and accurate matching results in Fig. 1 (d). Specifically, for matching cost calculation, the local feature vectors of rectified speckle images are extracted by census transform:

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

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

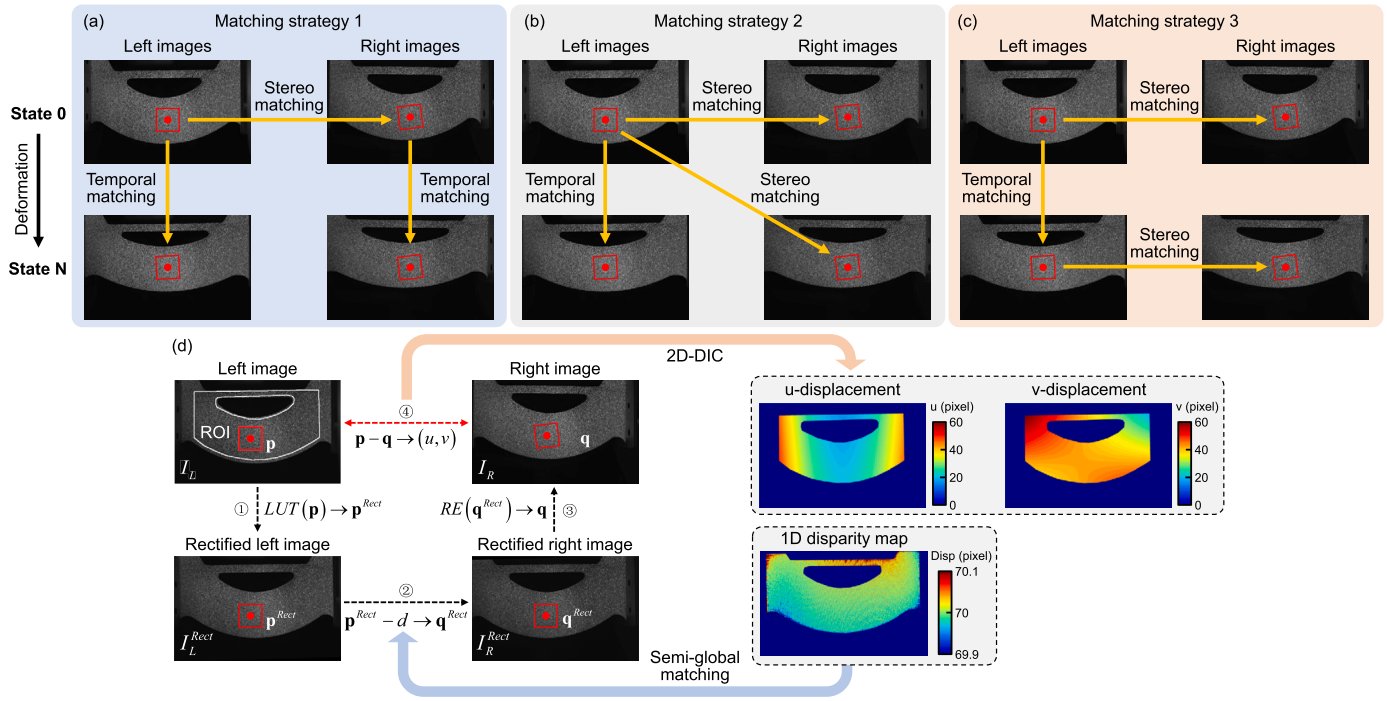

Fig. 1. Diagrams of different matching strategies for 3D deformation measurement and the proposed stereo matching method using semi-global matching with geometric constraints (GC-SGM). (a)-(c) Three different matching strategies, (d) An overview of the proposed GC-SGM.

where $C(x, y)$ is the feature vector of the central pixel (x, y), $\otimes$ represents a bit-wise concatenation operator, and R is the radius of the local matching window. Since the census transform based on block matching is a pixel-independent algorithm, the feature vectors of all pixels in stereo rectified images $\{I_L^{Rect}, I_R^{Rect}\}$ can be calculated simultaneously by multiple CUDA streams and the shared memory mechanism [16].

Based on the disparity range $[D_{min}, D_{max}]$ determined by the calibration parameters and measurement range of our Stereo-DIC system, the matching cost $Cost(x, y, d)$ 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:

$$Cost(x, y, d) = BC(C_L(x, y) \oplus C_R(x - d, y)), \tag{3}$$

where $\oplus$ is a XOR operation, and $BC(\bullet)$ is used to count the number of '1' in XOR results. Since the cost calculation using the Hamming distance is a row-independent algorithm, the initial matching cost of each row in the left rectified image can be calculated simultaneously benefiting from the shared memory mechanism. Due to the high overlap between candidate disparity ranges corresponding to adjacent pixels, the feature vectors of each row are stored in a segmented manner, and the matching costs of $D_{max} - D_{min} + 1$ pixels are computed at the same time.

And then, the matching cost $Cost(x, y, d)$ can be further optimized using cost aggregation or cost filtering. In our method, the SGM-based cost aggregation approximates the global solution by aggregating 1D matching costs along 4 independent paths. The aggregated cost $L_i(\mathbf{p}, d)$ of the pixel $\mathbf{p} = [x, y]^T$ at disparity d along a path $\mathbf{r}_i = (r_x, r_y)$ is defined recursively as:

$$L_i(\mathbf{p}, d) = Cost(\mathbf{p}, d) + \min \begin{cases} L_i(\mathbf{p} - \mathbf{r}_i, d) \\ L_i(\mathbf{p} - \mathbf{r}_i, d - 1) + P_1 \\ L_i(\mathbf{p} - \mathbf{r}_i, d + 1) + P_1 \\ \min\limits_k L_i(\mathbf{p} - \mathbf{r}_i, k) + P_2 \end{cases} - \min_k L_i(\mathbf{p} - \mathbf{r}_i, k), \tag{4}$$

$$S(\mathbf{p}, d) = \frac{1}{4} \sum_{i=1}^{4} L_i(\mathbf{p}, d), \tag{5}$$

where $S(\mathbf{p}, d)$ represents the aggregated cost. P_1 is a constant penalty, and P_2 is a penalty that varies with the intensity gradient:

$$P_1 \le P_2 = \frac{P_3}{\left| I(\mathbf{p}) - I\left(\mathbf{p} - \mathbf{r}_i\right) \right|} \le P_3, \tag{6}$$

where P_3 is another constant penalty. It is worth noting that the penalty parameters of SGM affect the final stereo matching results. After an exhaustive empirical search, it is found that the parameters P_1 and P_3 are closely related to the window radius R used in census transform. In order to enhance the accuracy of speckle matching, the preset thresholds of the parameters P_1 and P_3 are set as $(2R + 1)^2$ and $(4R + 2)^2$ based on the matching quality metric. Since the cost aggregation is a row-independent or column-independent algorithm, all four aggregated costs $L_i(\mathbf{p}, d)$ along different independent paths can be calculated simultaneously. Further, considering that the cost aggregation is performed sequentially along the row or column, the aggregated cost of the current pixel is only related to that of the previous pixel and can be calculated simultaneously for all candidate disparities.

Finally, the initial integer-pixel disparity map can be selected as the index of the minimum cost in $S(x, y, d)$ through Winner Take All (WTA):

$$D(x, y) = \underset{d}{\operatorname{argmin}}\, S(x, y, d). \tag{7}$$

Further, the sub-pixel disparity estimation is implemented by fitting a parabola using neighboring costs:

$$D^{sub} = D - \frac{S(D + 1) - S(D - 1)}{2S(D + 1) + 2S(D - 1) - 4S(D)}. \tag{8}$$

It is worth noting that the disparity computation and sub-pixel disparity estimation are pixel-independent algorithms, which are only related to the aggregated cost of the current pixel. Then some simple post-processing operations are adopted to obtain accurate and dense disparity maps. The left-right consistency check (L-R Check) is applied to identify invalid pixels in the disparity map, including occluded, mismatched, and background areas. The 4-connected-based image segmen-

tation algorithm is used for noise peak removal in the disparity map. The mismatched points in the disparity map are interpolated by choosing the second-lowest disparity value from the 8 neighbor points.

2.2. Initializing and accelerating Stereo-DIC computation using GC-SGM

After the 1D correspondence between the rectified stereo image is established using the optimized SGM based on GPU acceleration, the initial 2D displacement between raw stereo images can be found based on stereo calibration parameters and epipolar constraints. Using the intrinsic matrix $\mathbf{A}_L$ and the distortion coefficients $\mathbf{k}_L^c$ of the left camera, as shown in Fig. 1 (d), the point $\mathbf{p}$ located in the camera imaging plane I_L is mapped to the normalized plane, and further projected to the corresponding point $\mathbf{p}^{Rect}$ on the rectified left image I_L^{Rect} by combining rectified intrinsic parameters $\mathbf{A}_L^{Rect}$ and the rotation matrix $\mathbf{R}_L$ when performing the rectification of the left camera:

$$LUT(\mathbf{p}) \rightarrow \mathbf{p}^{Rect} \Leftrightarrow \mathbf{p}^{Rect} = \mathbf{A}_L^{Rect}\mathbf{R}_L^{-1}F_k^{-1}(\mathbf{k}_L^c, \mathbf{A}_L^{-1}\mathbf{p}), \tag{9}$$

where $F_k^{-1}(\bullet)$ represents the function of distortion removal. Eq. (9) describes the pixel-wise correspondence between the left image I_L and the rectified left image I_L^{Rect}. Since these parameters used in Eq. (9) are known and fixed after stereo calibration, a lookup table mapping from $\mathbf{p}$ to $\mathbf{p}^{Rect}$ can be created to calculate the sub-pixel coordinates of $\mathbf{p}^{Rect}$ directly. Then, the corresponding point $\mathbf{q}^{Rect}$ on the rectified right image I_R^{Rect} can be determined using the 1D disparity map D^{Rect} based on bicubic interpolation:

$$\mathbf{p}^{Rect} - d \rightarrow \mathbf{q}^{Rect} \Leftrightarrow \mathbf{q}^{Rect} = \mathbf{p}^{Rect} - D^{Rect}(\mathbf{p}^{Rect}). \tag{10}$$

Contrary to the lookup table mapping from $\mathbf{p}$ to $\mathbf{p}^{Rect}$, the point $\mathbf{q}$ on right image I_R can be converted from $\mathbf{q}^{Rect}$ according to the parameters of right camera:

$$RE(\mathbf{q}^{Rect}) \rightarrow \mathbf{q} \Leftrightarrow \mathbf{q} = \mathbf{A}_R F_k\left[\mathbf{k}_R^c, \mathbf{R}_R(\mathbf{A}_R^{Rect})^{-1}\mathbf{q}^{Rect}\right]. \tag{11}$$

It is worth noting that Eq. (11) is a pixel-independent algorithm, which is applicable for parallel acceleration. Here, the corresponding point $\mathbf{q}$ of right target image I_R can be determined for any point $\mathbf{p}$ on left reference image I_L, thus achieving accurate and efficient initial displacement estimation for stereo matching:

$$\mathbf{p} - \mathbf{q} \rightarrow (u, v) \Leftrightarrow \begin{cases} u = x_L - x_R \\ v = y_L - y_R \end{cases}. \tag{12}$$

For 2D-DIC-based sub-pixel refinement, the initial deformation parameters of the point $\mathbf{p}$ will be estimated using least-squares-based surface fitting. 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 the point $\mathbf{p}$ and the right target subset centered on the point $\mathbf{q}$ [13] as:

$$\sum_{\xi}(\mathbf{p}+\xi) = \sum_{\xi}(\mathbf{q} + \mathbf{W}(\xi;\mathbf{T})), \tag{13}$$

$$\mathbf{W}(\xi;\mathbf{T}) = \begin{bmatrix} x + u + u_x x + u_y y + u_{xy}xy + \frac{u_{xx}}{2}x^2 + \frac{u_{yy}}{2}y^2 \\ y + v + v_x x + v_y y + v_{xy}xy + \frac{v_{xx}}{2}x^2 + \frac{v_{yy}}{2}y^2 \end{bmatrix}, \tag{14}$$

where $\xi = [x, y]^T$ denotes the local coordinate in the subset, and $\mathbf{T}$ represents the deformation parameters, i.e., $[u, u_x, u_y, u_{xx}, u_{yy}, u_{xy}, v, v_x, v_y, v_{xx}, v_{yy}, v_{xy}]^T$. Using 2D matching results obtained according to Eq. (12), Eq. (13) can be rewritten based on the surface fitting:

$$S_x^2 = \sum_{\xi}(a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 - u)^2, \tag{15}$$

$$S_y^2 = \sum_{\xi}(b_1x^2 + b_2y^2 + b_3xy + b_4x + b_5y + b_6 - v)^2, \tag{16}$$

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. (15) and Eq. (16) 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 6 parameters without considering the second-order coefficients. Finally, sub-pixel optimization in stereo matching is performed using the GPU-accelerated 2D-DIC algorithm [12].

Table 1
Performance analysis results of Stereo-DIC using the SIFT-based method and GC-SGM.

Method	Computation time (ms)		Cmr[a]
	Initial estimation	Sub-pixel refinement	
SIFT	1411.60±17.31	**1691.48±30.73**	83.67%
GC-SGM	**134.41±8.52**	1710.64±34.35	**92.36%**

[a] Cmr = Correct matching rate (ZNCC ≥ 0.9).

3. Experiments

To systematically evaluate the effectiveness of the proposed GC-SGM, several experiments were conducted, including high-resolution 3D reconstruction of complex samples on the well-known 3D-DIC challenge dataset, 3D shape and deformation measurement on the built Stereo-DIC system. All tests were run on a desktop computer equipped with an AMD Ryzen 7 3700X CPU (8 cores, 16 threads) and an NVIDIA GeForce RTX3090 graphics card (10496 CUDA cores, 24 GB VRAM).

3.1. High-resolution 3D reconstruction on 3D-DIC challenge dataset

First, a 3D reconstruction experiment using Stereo-DIC was carried out to reveal the actual performance of GC-SGM in terms of the matching speed and accuracy, and the SIFT-based method [11] was implemented for comparison. As known to all, measuring objects with ridged and complex surfaces is a challenging task for Stereo-DIC. To verify the reliability of different methods for initial estimation in stereo matching, Stereo Sample 1 of the 3D-DIC dataset was measured, including two semi-cylindrical reliefs, two triangular prisms, and a square stage. Fig. 2 (a) shows the left image (2448 × 2048 pixels) of Stereo Sample 1 for 3D shape measurement, where the region of interest (ROI) was artificially set to 1561 × 1561 pixels with 1-pixel intervals. To maximize the matching performance, the subset size of 2D-DIC was determined as 19 × 19 pixels after an exhaustive empirical search. Figs. 2 (b)-(j) show the matching results combining the 2D-DIC algorithm with the SIFT-based method or GC-SGM, and the quantitative analysis results can be found in Table 1.

For initial 2D displacement estimation using the SIFT-based method, a large number of SIFT features were first extracted from left and right images, which are robust to deformation, distortion, and illumination changes. By calculating the Euclidean distance between SIFT descriptors, the features of the left image were paired with those of the right image. For each POI within the ROI, its initial deformation parameters were calculated using nearby feature pairs based on affine transformation [5]. Considering that rich features extracted from high-resolution speckle images are one of the main reasons affecting the running speed of SIFT-based methods [11], in this experiment, only 40,000 features were reserved to reach the balance between the matching efficiency and the correct matching rate. As shown in Figs. 2 (b)-(c), the SIFT-based method can cope well with flat regions, but is unable to achieve reliable initial 2D displacement estimation for slopes of triangular prisms and edges of cylinders, resulting in the valid points of the matching results being reduced from 1561 × 1561 pixels to 2,099,062 pixels (86.14%). After 2D-DIC-based sub-pixel refinement, high-quality matching points with ZNCC greater than 0.9 in the final u-/v-displacement results only account for 83.67% of all valid points in Table 1. The main reason for this result may be that the feature pairs to be matched in complex

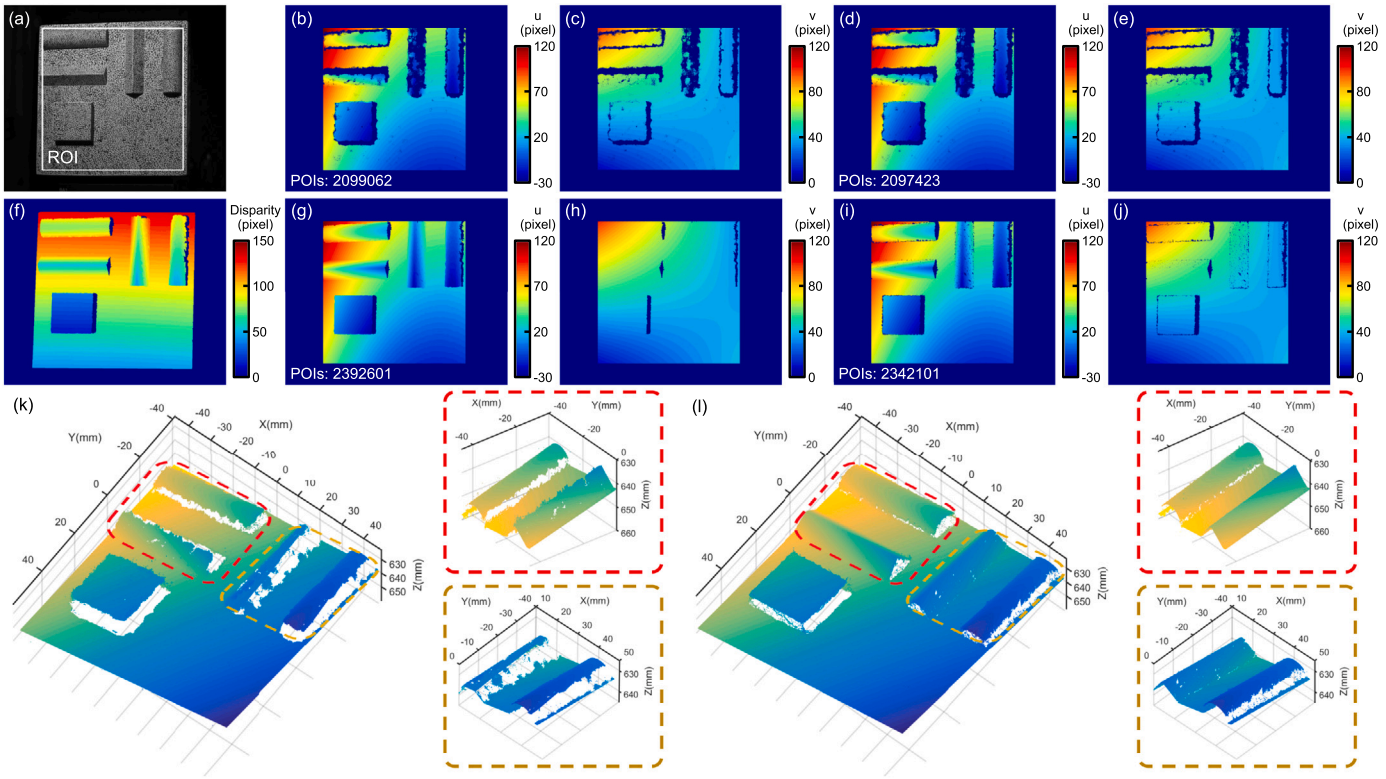

Fig. 2. Comparison of stereo matching results using different methods. (a) Left image of Stereo Sample 1 on 3D-DIC challenge dataset, (b)-(c) The initial u-/v-displacement results using SIFT-based method, (d)-(e) The refined u-/v-displacement results using 2D-DIC with SIFT-based method, (f) The 1D matching results using the GPU-accelerated SGM algorithm, (g)-(h) The initial u-/v-displacement results using the GPU-accelerated SGM algorithm with epipolar constraints, (i)-(j) The refined u-/v-displacements using 2D-DIC with GC-SGM, (k)-(l) 3D reconstruction results using SIFT-based method or GC-SGM.

surfaces have obvious location differences caused by perspective differences, which makes impossible the accurate initial estimation based on affine transformation. On the contrary, in the proposed GC-SGM, the GPU-accelerated SGM algorithm is utilized to achieve robust 1D matching of rectified stereo images via smoothing the disparity map, providing a dense 2D displacement result with 2,392,601 valid points (98.19%) as shown in Fig. 2 (g)-(h). Sub-pixel refinement using 2D-DIC with IC-GN was performed to further improve the correct matching rate to 92.36% in Figs. 2 (i)-(j). Based on stereo calibration parameters, 3D measurement results using different methods are presented in Figs. 2 (k)-(l), which confirm the superior performance of GC-SGM for high-quality and efficient 3D modeling of complex shapes.

For the computation time of different methods presented in Table 1, it can be found that SIFT still takes 1411.60±17.31 ms on Nvidia RTX3090 by parallel acceleration, due to the expensive computation involved in SIFT features extraction and matching. Benefiting from the optimized SGM based on GPU acceleration, robust and dense 1D matching takes about 120 ms for high-resolution images (2448 × 2048 pixels) with a disparity range of 160 pixels. Based on the 1D disparity map and stereo calibration parameters, geometric constraints and surface fitting are utilized to achieve pixel-independent 2D displacement estimation. The runtime of the whole initial estimation process requires only 134.41±8.52 ms, which is almost 10 times faster than SFIT-based methods [11]. In addition, GC-SGM needs to process more valid points in the 2D-DIC-based sub-pixel refinement stage in Fig. 2 (g)-(h). Since the initial matching accuracy of GC-SGM is higher than that of the SIFT-based method, especially in complex areas, so our method takes slightly more time to implement sub-pixel refinement than the SIFT-based method in Table 1. These experimental results demonstrate that our method achieves more accurate and computationally efficient 3D shape measurement based on Stereo-DIC.

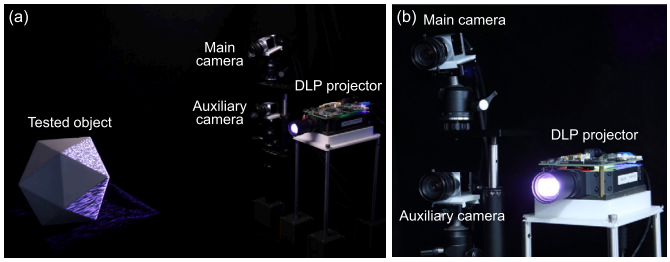

Fig. 3. Photographs of the built Stereo-DIC system including two CMOS cameras and a DLP projector.

3.2. 3D shape and deformation measurement on the built Stereo-DIC system

In order to verify the applicability of the proposed GC-SGM for 3D shape and deformation measurement, we built a common Stereo-DIC system with a wide baseline as shown in Fig. 3, which consists of two monochrome cameras (acA640-750um, Basler) with a resolution of 640 × 480 pixels and a DLP projector (LightCrafter 4500Pro, Texas Instruments). The projector projects the speckle pattern with a fixed spatial distribution onto the tested objects to ensure the global uniqueness of the entire measurement space and assist in establishing accurate correspondence between stereo images, enabling 3D shape measurement of texture-free surfaces [29]. To demonstrate the robustness of GC-SGM for measuring multiple objects with complex shapes, some different objects were measured in Figs. 4(a)–(c) including a hexagonal pyramid, a parallelepiped, and a regular icosahedron, and the corresponding 3D measurement results are presented to illustrate the reliability of our method as shown in Figs. 4(g)–(i). We additionally provide precision analysis results for 3D shape measurement of different objects. For the

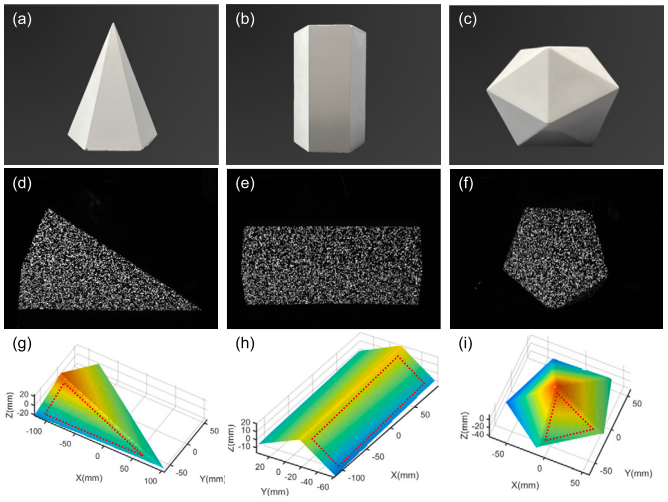

Fig. 4. 3D shape measurement of different objects. (a)-(i) The hexagonal pyramid, parallelepiped, and regular icosahedron, and the corresponding left speckle images and 3D reconstruction results.

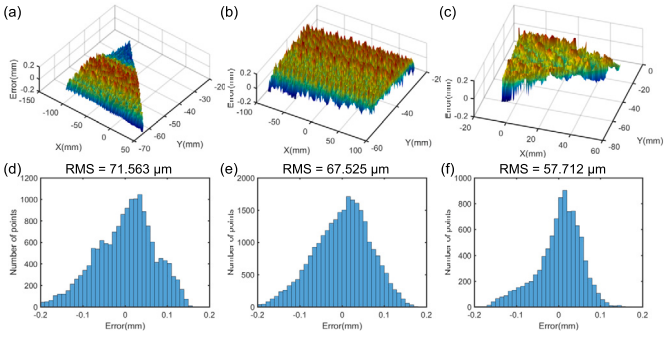

Fig. 5. Quantitative accuracy analysis for 3D shape measurement of different objects. (a)-(c) The error distributions of the hexagonal pyramid, parallelepiped, and regular icosahedron, (d)-(f) The corresponding histograms of (a)-(c).

highlighted regions of 3D shape measurement results in Figs. 4(g)–(i), the error distributions of the hexagonal pyramid, parallelepiped, and regular icosahedron are obtained by plane fitting as shown in Figs. 5 (a)-(c), where major measured errors are less than 200 μm with the RMS of 71.563 μm, 67.525 μm, and 57.712 μm in Figs. 5 (d)-(f). Quantitative analysis results for 3D shape measurement of different objects show that the proposed GC-SGM achieves accurate and robust 3D shape measurement for objects with complex surfaces, geometric steeps, and sharp edges.

Besides, in order to quantitatively evaluate the reliability of our method, the matching results of different measured objects were analyzed and shown in Table 2. The number of points is counted as the sum of all valid points from the main camera. To realize the reliability analysis of the matching results, the SIFT-based method and the proposed GC-SGM were performed as two separate groups. The correct matching rate represents the ratio of points with ZNCC greater than 0.9, and the time means the total running time of initial estimation and sub-pixel refinement. It can be found from these comparison results in Table 2 that GC-SGM realizes fast and robust 3D shape measurement with a high correctness ratio and high completeness for edge regions and abrupt depth discontinuities of complex objects. In addition, as the number of points increases from 60,068 POIs to 112,986 POIs, the computational time of SIFT improves significantly from 11.23 ms to 17.95 ms. Since the optimized SGM algorithm used in GC-SGM can quickly provide global matching results, the running time is still maintained

Table 2
Performance analysis results of several objects using SIFT-based method and GC-SGM.

Object	Nop[a]	Method	Cmr[b]	Time (ms)
Hexagonal pyramid	78651	SIFT	95.38%	13.04+28.63
		GC-SGM	**97.10%**	**4.96+29.01**
Parallelepiped	112986	SIFT	94.04%	17.95+41.21
		GC-SGM	**96.74%**	**5.14+42.14**
Regular icosahedron	60068	SIFT	92.93%	11.23+22.35
		GC-SGM	**95.58%**	**4.84+23.58**

[a] Nop = Number of points.
[b] Cmr = Correct matching rate (ZNCC ≥ 0.9).

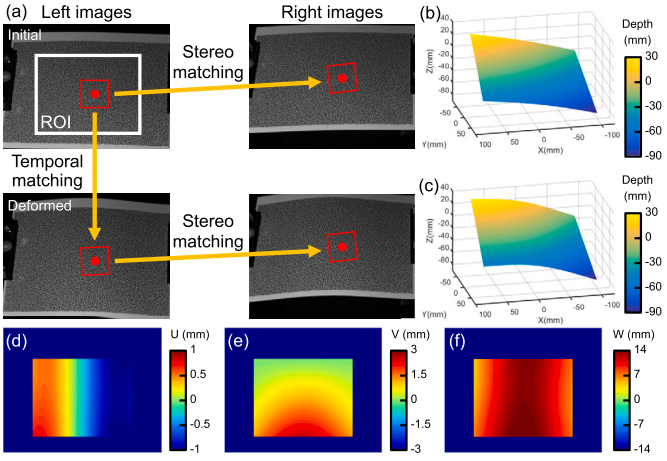

Fig. 6. 3D deformation results of the plastic plate. (a)-(c) Left and right images and the corresponding color-coded 3D shape reconstructions at different times, (d)-(f) U-/V-/W-displacement fields of 3D point clouds before and after deformation.

at about 5 ms for measuring these challenging shapes, demonstrating the reliability and efficiency of our method for 3D shape measurement based on Stereo-DIC.

Last, to further demonstrate the advantages of GC-SGM, our system was applied to fast 3D deformation measurement of the plastic plate as shown in Fig. 6 and Visualization 1. During the whole dynamic measurement, the plastic plate was subjected to external pressure from the left, and its shape gradually changed from flat to curved. For fast and accurate 3D deformation measurement, the matching window sizes of census transform and 2D-DIC are set as 5 × 5 pixels and 19 × 19 pixels, and the pixel spacing is 1 pixel. At the initial state, 3D shape measurement based on stereo matching was performed on the left and right images to obtain initial 3D point clouds as shown in Figs. 6(a)-(b). For subsequent deformation states, the SIFT-based method and GC-SGM are sequentially implemented for temporal matching and stereo matching. It is worth noting that the matched points of the deformed left image output by temporal matching are always at sub-pixel locations, as reference points for stereo matching. To circumvent the additional systematic errors caused by intensity interpolation, we referred to the adaptive subset offset scheme [15]. By transforming the subset in the reference image to the nearest integer position, this scheme not only retrenched the computational cost of subpixel interpolation, but also eliminated periodic systematic errors. Experiments prove that our system presents the 3D deformation measurement process of the plastic plate in quasi-real time (8.81 FPS) for calculating 120,701 POIs, which takes about 18.02+46.43 ms for temporal matching and 5.34+43.75 ms for stereo matching at each deformation state. These experimental results demonstrate that our method has finished renovating the

third matching strategy with poor performance, enabling accurate and computationally efficient 3D deformation measurement for objects undergoing large displacement or deformation.

4. Conclusions and discussion

In this work, we presented a reliable and efficient stereo matching method based on GC-SGM for initializing and accelerating Stereo-DIC computation, endowing the capability to overcome the limited matching efficiency and the reconstruction quality deterioration of complex profiles or large deformation in Stereo-DIC. An optimized semi-global matching (SGM) algorithm based on GPU acceleration is first presented to quickly establish global and dense 1D correspondence between rectified stereo images, which can be converted to initial 2D displacements between raw stereo images using epipolar constraints. For the second-order shape function involved in stereo matching, initial deformation parameters of each point are calculated using least-squares-based surface fitting, and then refined to enhance the accuracy of 3D shape and deformation measurement using 2D-DIC based on IC-GN. The effectiveness of GC-SGM has been verified by several experiments for measuring various types of objects. 3D reconstruction results of Stereo Sample 1 on 3D-DIC challenge dataset confirmed that GC-SGM can achieve high-precision and high-resolution 3D shape measurement for complex objects with steep and ridged surfaces, enhancing the matching correctness by about 10% while improving the processing speed on GPU by almost 10 times compared with SIFT-based methods. Based on the built Stereo-DIC system, 3D shape measurement results proved again that our method can achieve fast and robust 3D shape measurement with higher completeness for multiple objects including a hexagonal pyramid, a parallelepiped, and a regular icosahedron. Different from SIFT-based methods, the computational time of initial estimation based on GC-SGM is only related to the image resolution and is not affected by the number of points to be calculated, demonstrating the lower algorithmic complexity of our method for 3D shape measurement. Finally, dynamic measurement results of the plastic plate revealed the applicability of GC-SGM for accurate 3D deformation measurement with high quality in quasi-real time.

It should be discussed here that GC-SGM proposed in this paper can achieve more accurate and fast 3D measurement compared with SIFT-based methods, but it cannot meet the requirements of high-precision and high-resolution 3D measurement for complex structural parts in industrial scenarios. 2D-DIC-based sub-pixel optimization methods generally exploit larger matching windows to maintain the convergence stability and measurement robustness, but at the cost of the spatial resolution. In addition, how to select the appropriate matching window for Stereo-DIC is our primary focus. It is obvious that the smaller window can speed up the calculation efficiency but provides coarse correlation results. On the contrary, the bigger window outputs reliable corresponding points at the expense of the computational cost. Based on the above analysis, we will explore other methods to improve the matching accuracy and spatial resolution of Stereo-DIC.

CRediT authorship contribution statement

Wei Yin: Conceptualization, Methodology, Software, Visualization, Writing – original draft. **Yifan Ji:** Software, Data analysis. **Juntong Chen:** Data curation, Validation. **Rui Li:** Software, Data analysis. **Shijie Feng:** Writing – review & editing. **Qian Chen:** Formal analysis, Writing – review & editing, Funding acquisition. **Bing Pan:** Writing – review & editing. **Zhenyu Jiang:** Writing – review & editing. **Chao Zuo:** Formal analysis, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by National Key Research and Development Program of China (2022YFB2804603, 2022YFB2804604), National Natural Science Foundation of China (62075096, 62205147, 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), and National Major Scientific Instrument Development Project (62227818).

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.optlaseng.2023.107879.

References

[1] Sutton M, Hild F. Recent advances and perspectives in digital image correlation. Exp Mech 2015;55:1–8.
[2] Pan B. Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals. Meas Sci Technol 2018;29:082001.
[3] Chen D, Chiang F-P, Tan Y, Don H. Digital speckle-displacement measurement using a complex spectrum method. Appl Opt 1993;32:1839–49.
[4] Zhao J, Zeng P, Lei L, Ma Y. Initial guess by improved population-based intelligent algorithms for large inter-frame deformation measurement using digital image correlation. Opt Lasers Eng 2012;50:473–90.
[5] Zhou Y, Pan B, Chen YQ. Large deformation measurement using digital image correlation: a fully automated approach. Appl Opt 2012;51:7674–83.
[6] Wang Z, Vo M, Kieu H, Pan T. Automated fast initial guess in digital image correlation. Strain 2014;50:28–36.
[7] Bruck H, McNeill S, Sutton MA, Peters W. Digital image correlation using Newton-Raphson method of partial differential correction. Exp Mech 1989;29:261–7.
[8] Pan B, Li K, Tong W. Fast, robust and accurate digital image correlation calculation without redundant computations. Exp Mech 2013;53:1277–89.
[9] Shao X, Dai X, He X. Noise robustness and parallel computation of the inverse compositional Gauss–Newton algorithm in digital image correlation. Opt Lasers Eng 2015;71:9–19.
[10] Jiang Z, Kemao Q, Miao H, Yang J, Tang L. Path-independent digital image correlation with high accuracy, speed and robustness. Opt Lasers Eng 2015;65:93–102.
[11] Yang J, Huang J, Jiang Z, Dong S, Tang L, Liu Y, et al. Sift-aided path-independent digital image correlation accelerated by parallel computing. Opt Lasers Eng 2020;127:105964.
[12] Lin A, Li R, Jiang Z, Dong S, Liu Y, Liu Z, et al. Path independent stereo digital image correlation with high speed and analysis resolution. Opt Lasers Eng 2022;149:106812.
[13] Gao Y, Cheng T, Su Y, Xu X, Zhang Y, Zhang Q. High-efficiency and high-accuracy digital image correlation for three-dimensional measurement. Opt Lasers Eng 2015;65:73–80.
[14] Shao X, Dai X, Chen Z, He X. Real-time 3d digital image correlation method and its application in human pulse monitoring. Appl Opt 2016;55:696–704.
[15] Zhou Y, Sun C, Chen J. Adaptive subset offset for systematic error reduction in incremental digital image correlation. Opt Lasers Eng 2014;55:5–11.
[16] Yin W, Cao L, Zhao H, Hu Y, Feng S, Zhang X, et al. Real-time and accurate monocular 3d sensor using the reference plane calibration and an optimized sgm based on opencl acceleration. Opt Lasers Eng 2023;165:107536.
[17] Rishikesh K, Pramod R. Optical measurement techniques-A push for digitization. Opt Lasers Eng 2016;87:1–17.
[18] Zuo C, Qian J, Feng S, Yin W, et al. Deep learning in optical metrology: a review. Light: Sci Appl 2022;11:39.
[19] Yin W, Che Y, Li X, Li M, et al. Physics-informed deep learning for fringe pattern analysis. Opto-Electron Adv 2024;7:230034.
[20] Chang X, Bian L, Zhang J. Large-scale phase retrieval. eLight 2021;1:4.
[21] Luo Y, Zhao Y, Li J, et al. Computational imaging without a computer: seeing through random diffusers at the speed of light. eLight 2022;2:4.

[22] Qian J, Cao Y, Bi Y, et al. Structured illumination microscopy based on principal component analysis. eLight 2023;3:4.
[23] Wang L, Gong M, Gong M, Yang R. How far can we go with local optimization in real-time stereo matching. In: Third international symposium on 3D data processing, visualization, and transmission. IEEE; 2006. p. 129–36.
[24] Gallup D, Frahm J-M, Mordohai P, Yang Q, Pollefeys M. Real-time plane-sweeping stereo with multiple sweeping directions. In: 2007 IEEE conference on computer vision and pattern recognition. IEEE; 2007. p. 1–8.
[25] Woetzel J, Koch R. Real-time multi-stereo depth estimation on gpu with approximative discontinuity handling. In: 1st European conference on visual media production; 2004.
[26] Sun J, Zheng N-N, Shum H-Y. Stereo matching using belief propagation. IEEE Trans Pattern Anal Mach Intell 2003;25:787–800.
[27] Hirschmüller H. Stereo processing by semiglobal matching and mutual information. IEEE Trans Pattern Anal Mach Intell 2007;30:328–41.
[28] Hirschmüller H, Scharstein D. Evaluation of stereo matching costs on images with radiometric differences. IEEE Trans Pattern Anal Mach Intell 2008;31:1582–99.
[29] Yin W, Hu Y, Feng S, Huang L, Kemao Q, Chen Q, et al. Single-shot 3d shape measurement using an end-to-end stereo matching network for speckle projection profilometry. Opt Express 2021;29:13388–407.