Research Article | Vol. 28, No. 8 / 13 April 2020 / Optics Express 12219

Optics EXPRESS

# Geometry constrained correlation adjustment for stereo reconstruction in 3D optical deformation measurements

ZHILONG SU,[1,2] LEI LU,[3] FUJUN YANG,[4] XIAOYUAN HE,[4,5] AND DONGSHENG ZHANG[1,2,6]

[1] *Shanghai Institute of Applied Mathematics and Mechanics, School of Mechanics and Engineering Science, Shanghai University, Shanghai 200444, China*
[2] *Shanghai Key Laboratory of Mechanics in Energy Engineering, Shanghai 200444, China*
[3] *College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China*
[4] *School of Civil Engineering, Southeast University, Nanjing 210096, China*
[5] *mmhxy@seu.edu.cn*
[6] *donzhang@staff.shu.edu.cn*

**Abstract:** Recovering the geometric shape of deformable objects from images is essential to optical three-dimensional (3D) deformation measurements and is also actively pursued by researchers. Most of the existing techniques retrieve the shape data with triangulation based on pre-estimated stereo correspondences. In this paper, we instead propose to recover depth information directly from images of a binocular vision system for 3D deformation estimation. Given a calibrated geometry of the system, the reprojection error is parameterized by the depth and then described with local intensity dissimilarity between a stereo pair in considering spatial deformation. Afterward, a correlation adjustment model is formulated to estimate the depth parameter by minimizing the error. As a solving strategy, we show the Gauss-Newton linearization of the proposed model and its initialization. 3D displacement estimation based on depth information is also presented. Experiments, including rigid translation and bending deformation measurements, are conducted to verify the performance of the proposed method. Results show that the proposed method is accurate yet precise in 3D deformation estimations. Other underlying developments are underway.



## 1. Introduction

Recent years have seen significant progress on the problem of measuring deformation with digital image correlation (DIC) techniques. Measuring the surface deformation of objects has been significantly altered through the adoption of subset-based DIC algorithms. Among them, the inverse compositional Gauss-Newton (IC-GN) algorithm has attracted much attention owing to the high-efficiency in computation. Since the algorithm was proposed in image registration [1], researchers have been sparing no efforts in improving the accuracy [2–4], robustness to image noise and intensity variations, as well as efficiency [5–7]. This has led to steady but incremental progress on optical deformation measurement approaches. Because of the outstanding performance of IC-GN, it also provides an efficient way to process the perspective deformation to this study.

The combination of DIC and stereophotogrammetry has spawned a valuable technique on three-dimensional (3D) deformation measurements, the stereo digital image correlation (stereo-DIC) [8]. It has proved to have attractive abilities in measuring shape and deformation of arbitrary surfaces. As the technique starts to show advantages such as non-contact and full-field measurements, it is heralding a revolution in 3D deformation data acquisition in contrast to the

#392248
Journal © 2020

https://doi.org/10.1364/OE.392248
Received 9 Mar 2020; revised 3 Apr 2020; accepted 3 Apr 2020; published 8 Apr 2020

traditional methods (such as strain gauges, mechanical or laser extensometers). As a result, huge parts of daily requirements in experimental mechanics, e.g., strain analysis of biological tissues [9], human body [10] and transparent materials [11], deformation monitoring and control in industrial manufacturing [12–14], large-scale engineering structure measurements [15,16], and many more are powered by stereo-DIC. In order to produce significant results in 3D deformation measurements, researchers have not only played to the strength of the fundamental binocular imaging systems, but also developed multi-camera measurement systems [17,18], and even devised several compact stereo imaging devices based on single camera [19–21]. The constant advances in stereo imaging techniques and measurement systems allow for fine displacement and strain studies of materials to screen for ideal experimental results.

Despite the above different 3D optical deformation measurement systems and applications that can be found, the fundamental problems to be addressed are often similar, mainly including stereo camera calibration, temporal-matching, and 3D reconstruction. Stereo calibration aims to determine the internal parameters of and the external geometry between cameras with artificial target-based techniques [22,23] or auto-calibration methods [24,25]. Temporal-matching is applied to track the deformed positions of points in an image sequence. Owning to intensive studies on the IC-GN based subset matching [1,2,5,7], the accuracy of temporal-matching was reported up to 0.01 pixels when following well-controlled experiment conditions [3,26]. On the basis of the previous two, the 3D reconstruction is conducted to obtain the shape information of objects for computing 3D displacement data.

Traditionally, the shape of deformable objects is recovered from the point correspondences built across two calibrated cameras. The process often follows the classic pipeline of firstly establishing correspondences with stereo-matching followed by reconstructing the 3D profile through triangulation [8]. However, the reconstruction procedure does not utilize the image information again, leading to the accuracy and precision of 3D shape reconstruction are heavily dependent on the quality of the calibration and stereo correspondences. While several studies investigated how the accuracy of stereo-DIC is affected by errors in 3D reconstruction from different aspects, e.g., camera calibration [27,28], stereo-matching [29,30] and camera self-heating [31], few attentions are paid to the improvement of 3D shape reconstruction method for optical deformation measurements.

Therefore, an end-to-end framework based on stereo vision geometry and correlation is proposed to obtain high-quality depth reconstruction for 3D displacement estimation. The model not only works without any pre-estimated stereo correspondences but also improves the measurement accuracy and robustness by directly retrieving 3D information in the image domain. In a given binocular stereo imaging setup, the back-projection is firstly formulated with a depth parameterization. Then a reprojection is obtained under the vision geometry constraint. The error of the reprojection is expressed in terms of intensity discrepancies between a stereo pair, giving a direct connection between the depth and image information. This allows us to directly recover the depth by solving a stereo image correlation problem. We propose the method as geometry constrained correlation adjustment, and refer to as GC-CA for short. Because of perspective deformation, the GC-CA model is built in the form of inverse compositional correlation, maintaining a high-efficiency in computation yet improving the robustness to image noise. Gauss-Newton based solving approach, as well as its initialization strategies and implementation details, are provided for the proposed model to enhance the feasibility and quality of the depth estimation. With the reconstructed depth information, the 3D displacement estimation algorithm is presented finally. To our knowledge, few studies in deformation measurements have looked at the issue of estimating 3D shape directly from the calibrated images. Our study fills in the research gap by providing the proposed stereo depth reconstruction framework.

The rest of this pater is organized as follows: Section 2 presents the GC-CA model. The principle of the proposed framework is established in Section 2.1. The Gauss-Newton optimization

Research Article Vol. 28, No. 8 / 13 April 2020 / Optics Express 12221


algorithm is described in Section 2.2. The initialization strategies and implementation details are given in Sections 2.3 and 2.4, respectively. The process of 3D displacement estimation is presented in Section 3. Section 4 gives the experimental verifications. The accuracy of the proposed model is investigated in Section 4.1, and the performance in 3D deformation measurement is studied in Section 4.2. Section 5 concludes.

## 2. Geometry constrained correlation adjustment framework

In this section, the GC-CA framework, which reformulates the IC-GN algorithm based on the constraint of the calibrated vision geometry, is presented. Subsequently, the Gauss-Newton based optimization algorithm and the depth initialization are introduced. We end with a description of the implementation details.

### 2.1. Principle of GC-CA

Given that a 3D optical deformation measurement system composed of two cameras, the schematic geometry and principle overview of the proposed framework are shown in Fig. 1. For the sake of conciseness, we assume that the imaging model is ideal and thus image distortion is not considered temporarily. The explanation of GC-CA is also based on the simple sum of squared differences (SSD) criterion, but the same derivations stay valid for any other correlation criteria.

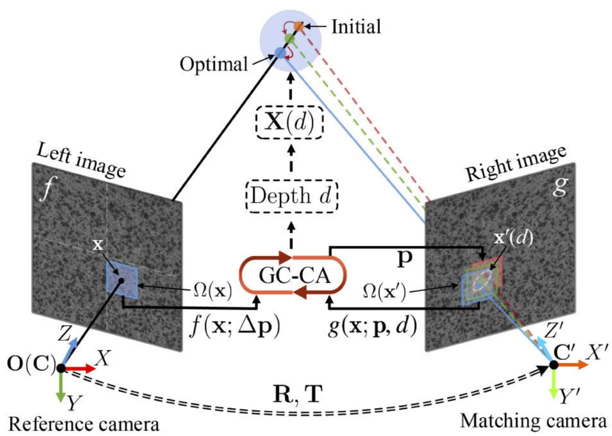


**Fig. 1.** The geometry and schematic illustration of the GC-CA framework, where the object coordinate frame X-Y-Z is aligned and attached to the reference frame **C**. Because of epipolar geometry constraint, the projection $\mathbf{x}'(d) \in \mathbb{R}^2$ of object point $\mathbf{X}(d) \in \mathbb{R}^3$ produced by adjusting the depth $d$ should move along a straight line; meanwhile, the deformation parameter vector $\mathbf{p}$ is also updated to warp the subset $\Omega(\mathbf{x}')$ surrounds $\mathbf{x}'(d)$ to match against the one $\Omega(\mathbf{x})$ centered at $\mathbf{x}$.

Considering that the camera system is well pre-calibrated using stereo calibration techniques such as [22,25], a fixed, precise external imaging geometry can be obtained. We follow a reasonable assumption that the object coordinate system is aligned to the frame of the left camera **C**. The camera is accordingly referred to as the reference and the right $\mathbf{C}'$ is defined as the matching camera in this paper. With the frame configuration, an object point can be parameterized by its depth $d$ along the light ray passing through its projection $\mathbf{x} \in \mathbb{R}^2$ in the left image $f$ and the optical center of the reference camera. According to the pinhole camera model, the counterpart $\boldsymbol{x} \in \mathbb{R}^3$ of $\mathbf{x}$ in the normalized image domain can be obtained with the following

Optics EXPRESS

inverse mapping:

$$\boldsymbol{x} = \mathbf{K}^{-1} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \tag{1}$$

where

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times3}$$

is the pre-calibrated intrinsic matrix of the reference camera, with $f_x$ and $f_y$ being the focal lengths in pixel dimensions in $x$- and $y$-directions, and $(c_x, c_y)$ being the principle point in pixel dimensions.

Relative to the reference frame X-Y-Z shown in Fig. 1, the object point back-projected from $\mathbf{x}$ can be expressed by scaling the normalized term $\boldsymbol{x}$ in Eq. (1) using its depth $d$ as:

$$\mathbf{X}(d) = \boldsymbol{x}d. \tag{2}$$

For the calibrated stereo imaging system, the back-projection above defines the 3D point in form of depth parametrization, so that $\mathbf{X}(d)$ can be identified in terms of its projection in the reference camera up to the only unknown parameter, $d$. This allows reducing the degree of freedom for the problem of 3D point reconstruction from 3 to 1. As the depth is determined, the object point is then recovered.

In order to estimate depth $d$, the reprojection of the object point $\mathbf{X}(d)$ on the image plane of the matching camera is considered. Let $\mathbf{R} \in SO(3)$ be the pre-calibrated relative rotation matrix and $\mathbf{T} \in \mathbb{R}^3$ be the relative translation vector from the reference frame to the matching camera frame. As shown in Fig. 1, the projection $\mathbf{x}'(d)$ of $\mathbf{X}(d)$ in the right image $g$ is thus given by:

$$[\mathbf{x}'(d), 1]^T = \mathbf{K}'\langle \mathbf{R}\mathbf{X}(d) + \mathbf{T}\rangle, \tag{3}$$

where $\mathbf{K}'$ denotes the known intrinsic matrix of the matching camera, and operator $\langle\cdot\rangle$ indicates the normalization that maps $[x, y, z]^T$ to $[x/z, y/z, 1]^T$. Clearly, $\mathbf{x}'(d)$ is also the function of $d$. At this point, it is ready to identify the final depth.

Conventionally, the depth could be determined by trivially minimizing the Euclidean distance between the reprojection $\mathbf{x}'(d)$ and the observed correspondence (obtained by pre-matching with $\mathbf{x}$). However, using pre-matched point correspondence to reconstruct depth does not use image information again, posing that the quality of 3D reconstruction heavily depends on the accuracy of the stereo matching. One possible way of addressing the problem is to bring the image back into the loop of stereo reconstruction. Drawing inspiration from the traditional subset matching, we devise the GC-CA framework to identify the depth parameter $d$ from the stereo pair $f$ and $g$ directly. As shown in Fig. 1 and Eq. (3), each of $\mathbf{x}'(d)$ is obtained based on the given geometry and thus gives rise to an implicit constraint to enable a feasible correlation adjustment. Because of the perspective deformation, the inverse compositional approach to the local spatial deformation is adopted below in modeling the correlation adjustment framework. This enables an end-to-end stereo depth reconstruction with high-accuracy for 3D displacement computation.

Let $\Omega = \{\boldsymbol{\eta} = (\eta_x, \eta_y) | -M \le \eta_x, \eta_y \le M\}$ be a square window with a radius of $M$, the subset surrounds $\mathbf{x}'(d)$ can be described as $\Omega(\mathbf{x}') = \{g(\mathbf{x}'(d) + \boldsymbol{\eta}) | \boldsymbol{\eta} \in \Omega\}$, where $d$ is omitted in $\Omega(\mathbf{x}')$ for notation clarity. Similarly, the subset centered at the point $\mathbf{x}$ is given by $\Omega(\mathbf{x}) = \{f(\mathbf{x} + \boldsymbol{\eta}) | \boldsymbol{\eta} \in \Omega\}$. The fundamental model of the proposed GC-CA framework can be established using the inverse

Optics EXPRESS

compositional SSD as follows:

$$C(d, \mathbf{p}) = \frac{1}{2} \sum_{\boldsymbol{\eta} \in \Omega} [f(\mathbf{x} + \mathcal{T}(\boldsymbol{\eta}; \Delta \mathbf{p})) - g(\mathbf{x}'(d) + \mathcal{T}(\boldsymbol{\eta}; \mathbf{p}))]^2 , \tag{4}$$

where $\mathbf{p} = [u_x, u_y, v_x, v_y]^T$ is the deformation parameter vector of the subset $\Omega(\mathbf{x}')$, $\Delta\mathbf{p}$ indicates the incremental vector of $\mathbf{p}$, and

$$\mathcal{T}(\boldsymbol{\eta}; \mathbf{p}) = \begin{bmatrix} 1 + u_x & u_y \\ v_x & 1 + v_y \end{bmatrix} \begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix}, \tag{5}$$

denotes the first-order spatial shape transform of the subset $\Omega(\mathbf{x}')$ relative to its center. The shape transformer $\mathcal{T}(\boldsymbol{\eta}; \mathbf{p})$ deforms the shape of the subset with the given linear transformation, endowing a capacity to the model to cope with the projective deformation. It should be noted that any differentiable spatial transformers could be used in the model if needed. For general 3D surfaces, a nonlinear spatial transformer formulated with the first five rows and columns of the second-order shape function [6] is recommended for gaining better matching to the local projective deformation. Finally, the depth $d$ can be computed by solving the nonlinear problem above.

In contrast to the traditional stereo reconstruction accomplished with an explicit stereo-matching followed by the triangulation, the proposed model directly outputs the depth information by invoking a correlation adjustment when the first reprojection is done. In the process, as the subset $\Omega(\mathbf{x}')$ is gradually warped and moved to the optimal matching position, the depth $d$ is gradually updated to the optimal one. For ease of understanding, the procedure is schematically drawn as the dashed lines in Fig. 1. It is worth mentioning that, because the vision geometry is mainly supported by the calibration parameters, the error propagation due to the imprecise calibration of the stereo imaging system results in systematic errors . Moreover, we choose the inverse compositional version versus the classic forwards additive counterpart because the efficiency of the depth refinement can be improved significantly by maintaining most terms stay in constant over Gauss-Newton iterations. This is described in detail in Section 2.2.

### 2.2. Gauss-Newton optimization

Equation (4) models the depth estimation as a nonlinear correlation problem. By minimizing the objective function, the optimal depth can be determined. There are various standard methods to accomplish this. Because of the inverse compositional form of the proposed framework, an efficient yet robust way is to use the Gauss-Newton algorithm. However, the presence of the depth parameter brings several differences in the aspects of linearization, gradient computation, and parameter updating, compared with the traditional IC-GN algorithm. Therefore, the Gauss-Newton optimization for the proposed GC-CA is presented.

By applying the Taylor expansion theorem in Eq. (4), the linearized objective function is obtained by truncating the higher-order terms as follows:

$$L(\Delta d, \Delta \mathbf{p}) = \frac{1}{2} \sum_{\boldsymbol{\eta} \in \Omega} \left[ \epsilon(d, \mathbf{p}) + \frac{\partial \epsilon(d, \mathbf{p})}{\partial d} \Delta d + \frac{\partial \epsilon(d, \mathbf{p})}{\partial \mathbf{p}} \Delta \mathbf{p} \right]^2 , \tag{6}$$

where

$$\epsilon(d, \mathbf{p}) = f(\mathbf{x} + \boldsymbol{\eta}) - g(\mathbf{x}'(d) + \mathcal{T}(\boldsymbol{\eta}; \mathbf{p})) \tag{7}$$

is pixel-wise intensity residual between the subsets $\Omega(\mathbf{x})$ and $\Omega(\mathbf{x}')$,

$$\frac{\partial \epsilon(d, \mathbf{p})}{\partial d} = -\nabla g \frac{\partial \mathbf{x}'}{\partial d} \tag{8}$$

Optics EXPRESS

and

$$\frac{\partial \epsilon(d,\mathbf{p})}{\partial \mathbf{p}} = \nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}} \tag{9}$$

are the derivatives of the intensity residual with respect to the depth and spatial deformation parameters, respectively. $\nabla f$ and $\nabla g$ are the gradients of images $f$ and $g$ at points $\mathbf{x}$ and $\mathbf{x}'(d) + \mathcal{T}(\boldsymbol{\eta};\mathbf{p})$, respectively. $\frac{\partial \mathbf{x}'}{\partial d}$ is the derivative of the reprojected point $\mathbf{x}'(d)$ with respect to the depth parameter, and

$$\frac{\partial \mathcal{T}}{\partial \mathbf{p}} = \begin{bmatrix} \eta_x & \eta_y & 0 & 0 \\ 0 & 0 & \eta_x & \eta_y \end{bmatrix}$$

is the Jacobian of the spatial transformer.

Equation (6) is a function of the increments of the depth and deformation parameters, $\Delta d$ and $\Delta\mathbf{p}$. The goal of the Gauss-Newton optimization is to compute both of the increments for updating $d$ and $\mathbf{p}$ to a new state. For this purpose, taking derivatives of the function in Eq. (6) with respect to the increments $\Delta d$ and $\Delta\mathbf{p}$ respectively, then let them be zeros, we obtain the normal equations of the Gauss-Newton algorithm as:

$$\mathbf{H}\begin{bmatrix} \Delta d \\ \Delta \mathbf{p} \end{bmatrix} = \sum_{\boldsymbol{\eta}\in\Omega} \begin{bmatrix} -\nabla g \frac{\partial \mathbf{x}'}{\partial d} \\ \left(\nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}}\right)^T \end{bmatrix} \epsilon, \tag{10}$$

where $\mathbf{H}$, the Gauss-Newton approximation of the Hessian with dimensions of $5 \times 5$, is given by:

$$\mathbf{H} = \sum_{\boldsymbol{\eta}\in\Omega} \begin{bmatrix} \left(\nabla g \frac{\partial \mathbf{x}'}{\partial d}\right)^2 & -\nabla g \frac{\partial \mathbf{x}'}{\partial d} \nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}} \\ -\nabla g \frac{\partial \mathbf{x}'}{\partial d} \left(\nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}}\right)^T & \left(\nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}}\right)^T \nabla f \frac{\partial \mathcal{T}}{\partial \mathbf{p}} \end{bmatrix}. \tag{11}$$

After $\Delta d$ and $\Delta\mathbf{p}$ are solved from Eq. (6), the depth value and the deformation vector can be updated by the following equations:

$$\begin{cases} d \leftarrow d + \Delta d \\ \mathcal{T}(\mathbf{p}) \leftarrow \mathcal{T}(\mathbf{p}) \circ \mathcal{T}^{-1}(\Delta \mathbf{p}) \end{cases} \tag{12}$$

where $\circ$ is the compositional operator [1]. Equations (10)–(12) show the iterative format of the Gauss-Newton algorithm for the proposed GC-CA framework. In addition, Eq. (11) shows an essential requirement for solving the problem is that the intensity gradient response within the subsets should not be vanishing, and should above the image noise level. This condition is usually satisfied with speckle patterns but is not always be satisfied when using natural surface textures. For that, it is recommended to estimate the depths by extracting salient feature points in the natural surface patterns [32].

For the Gauss-Newton in Eq. (10), it is worth noting that the gradient contributions in terms of the deformation parameters, $\frac{\partial \epsilon}{\partial \mathbf{p}}$, are invariant over iterations because $\nabla f$ and $\frac{\partial \mathcal{T}}{\partial \mathbf{p}}$ in Eq. (9) are both evaluated on the subset $\Omega(\mathbf{x})$. This means that the entries, corresponding to the deformation parameters, in the Hessian matrix and the right-hand side vector remain constant over iterations. In Eq. (10), it clearly shows that these constant terms in iteration can be grouped together to obtain a constancy pattern: the lower right $4 \times 4$ sub-matrix in $\mathbf{H}$, and the last 4 elements in the right-hand side vector are constant. The constancy pattern gained from the inverse compositional form greatly reduces the computational load due to the majority of gradient evaluations are avoided, ensuring the efficiency of the proposed GC-CA. This advantage will be more significant when a higher-order spatial transform pattern is used in Eq. (5).

Optics EXPRESS

### 2.3. Initialization

Initialization is also necessary to depth estimation in the GC-CA framework. A good initialization benefits the performance in aspects of convergence and speed. For each point computation, the initial guess to the deformation parameter vector $\mathbf{p}$ is straightforward by setting $\mathbf{p} = \mathbf{0}$ at the onset of computation. However, the depth initialization is not so trivial compared with that for $\mathbf{p}$, because there is no direct apriori information for the depth to be recovered. To address the problem, a reasonable approach to provide the depth initial guess is to retrieve some information on the depth that is available in pre-calibration. Since control points used in the pre-calibration are often close to or on the surface of the object being measured, the range of depth variations between the control points and the object surface is quite small relative to the object distance. Accordingly, the depth information of each control point can be used to initialize the depth reconstruction in the undeformed configuration. We refer to these known depths in the pre-calibration as nominal depths to the depth initialization.

With the nominal depths, a closest depth assignment (CDA) strategy is advocated to improve the quality of the depth initialization. For one initial object point $\mathbf{X}(d)$ to be reconstructed, its nearest control point is searched by testing the distance between the projection of $\mathbf{X}(d)$ and the projections of all control points observed from the reference camera. Then the nominal depth corresponding to the searched control point is used as the initial guess in computing $d$. It is worth noting that, if the artificial target-based calibration is employed, the control points are suggested to be selected from the calibration pose placed in the middle of the field of view and closest to the object, whenever it is possible. For the self-calibration techniques where the control points come from surfaces being measured, the pre-determined depth of each control point can be directly used as a nominal depth. After the depth map in the initial state is estimated, one can initialize subsequently depth estimation using the result of its previous state or the initial depth estimation according to the magnitude of out-of-plane deformation. This is feasible since the depth variations caused by the object deformation is rather small, not least for the widely used solid materials and structures. We found that the CDA strategy could ensure the subset $\Omega(\mathbf{x}')$ mostly overlaps with the optimal one corresponding to the subset $\Omega(\mathbf{x}')$, so that the problem in Eq. (4) can be solved effectively by the above Gauss-Newton algorithm.

### 2.4. Implementation details

For notation brevity, the GC-CA model is established based on the SSD criterion and the ideal pinhole model in Section 2. Here several implementation details are introduced to enhance the practicability of the proposed method. In comparison to the SSD, its robust variant, zero-mean normalized SSD (ZNSSD), is more practical [5], and thus is recommended to model the problem in Eq. (4).

Moreover, because of the imperfect optical systems of cameras in practice, the lens distortion should be considered in the proposed framework, not least that the radial distortion. The radial distortion models introduced in literatures [25] and [22] are recommended for the back-projection and the reprojection in Eqs. (2) and (3), respectively. Although both distortion functions can be expressed in the same symbolic form, they behave in opposite ways: the former maps a distorted image point to its undistorted counterpart analytically, and the latter distorts an ideal image point. This implies that both the back-projection and the reprojection can maintain an exact analytic form, rather than the iterative procedure used in the tradition. By using this radial distortion process scheme, one significant benefit is the derivate term $\frac{\partial \mathbf{x}'}{\partial d}$ can be computed analytically. The other benefit is that the accuracy loss caused by the distortion rectification is mitigated. Both are essential for the Gauss-Newton optimization in Section 2.2. Note that the distortion functions of the reference and matching cameras should be determined respectively in the system pre-calibration.

Optics EXPRESS

## 3. 3D displacement field estimation

The estimation of 3D displacement is implemented by comparing the deformed object shape with its initial counterpart (at the undeformed stage). Let $\mathbf{x}_0$ be the initial position of the current point $\mathbf{x}$ imaged by the reference camera. Given the depths estimated by the proposed GC-CA corresponding to $\mathbf{x}_0$ and $\mathbf{x}$ are $d_0$ and $d$, respectively, the corresponding object points $\mathbf{X}(d_0)$ and $\mathbf{X}(d)$ are obtained according to Eq. (2). By substituting Eq. (1) into Eq. (2), the 3D displacement vector can be estimated in terms of the depth as:

$$\mathbf{U} = \mathbf{X}(d) - \mathbf{X}(d_0) = \mathbf{K}^{-1} \begin{bmatrix} \Delta d\mathbf{x}_0 + d\mathbf{u} \\ \Delta d \end{bmatrix} \tag{13}$$

where $\Delta d = d - d_0$ is the depth variation, $\mathbf{u} \in \mathbb{R}^2$, the displacement from image point $\mathbf{x}_0$ to $\mathbf{x}$, is determined by the temporal-matching.

According to Eq. (13), the overall pipeline of the 3D displacement computation consists of pre-calibration, initial depth evaluation, temporal image matching, depth reconstruction, and displacement output, which is illustrated in Fig. 2. In summary, the overall computing procedure is listed below:

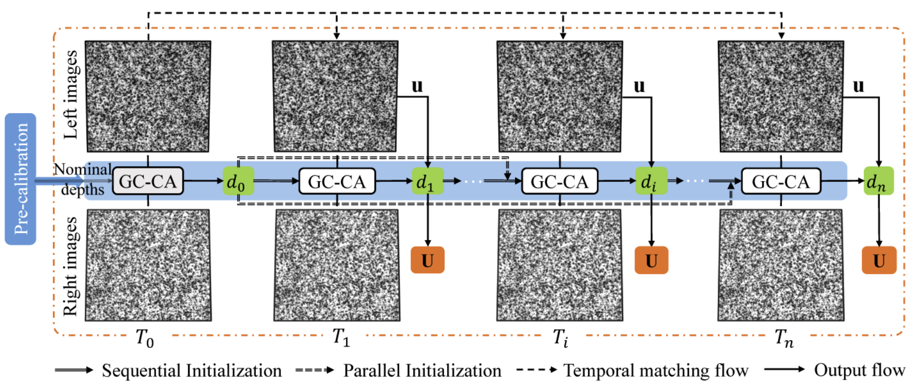


**Fig. 2.** 3D displacement estimation pipeline with the proposed GC-CA algorithm. There are two feasible initialization flows for depth computing presented: sequential and parallel initialization. The former allowed for cases with large out-of-plane deformation because the initial guess comes from the previous deformation state, while the latter is suitable for small out-of-plane deformation since the depth estimation after deformation uses the depth information evaluated at begin.

Step 1: Pre-calibration of the 3D optical measurement system. The intrinsic matrices, $\mathbf{K}$ and $\mathbf{K}'$, for both cameras as well as the external parameters, $\mathbf{R}$ and $\mathbf{T}$, are determined in advance. In addition, a set of nominal depths can be obtained for the CDA initialization according to the calibrated geometry between the reference camera and the calibration target. In this sense, the stereo calibration methods [24,25] are recommended to obtain better depth initialization.

Step 2: Initial depth evaluation. In full-field optical measurements, obtaining deformation data for densely distributed points of interest (POIs) is desired. For that, this step aims to recover the initial depth $d_0$ for every POI in the undeformed state $T_0$ using the proposed GC-CA. The initial guess is obtained with the CDA strategy, as described in Section 2.3. We propose this strategy following the continuity hypothesis in 3D deformation measurements that is, for two points on the object surface being measured, the closer their projections the smaller difference in their depths.

Step 3: Temporal image matching. This step computes the deformed position $\mathbf{x}$ and the displacement vector $\mathbf{u}$ for each of POIs. They can be obtained using 2D image registration

Optics EXPRESS

techniques. Since the contribution of **u** to the 3D displacement is scaled by the depth $d$, it is required to estimate **u** as accurately as possible. Hence, the IC-GN based DIC method is highly recommended in this step. Note that there is no order of priorities between this and the previous steps. They can be performed simultaneously.

Step 4: Depth reconstruction after deformation. For each deformed state $T_i(i = 1, \ldots, n)$, the depth $d_i$ corresponding to **x** is recovered from the current image pair using the proposed GC-CA. Two possible ways are provided to obtain the initial guess in this step, as illustrated in Fig. 2. One of them is performed sequentially, that is, the initial value of the current depth $d_i$ is the corresponding value $d_{i-1}$ in the previous state; the other is the initial guess of $d_i$ comes from the corresponding depth $d_0$ in the initial state, enabling the depth reconstruction can be performed in parallel.

Step 5: Calculates 3D displacements, by substituting the results in the above steps into Eq. (13).

## 4. Experimental verifications

The proposed algorithm was verified in the section to show its correctness and performance in displacement estimation and strain measurement, by using a rigid body translation and plate bending experiments respectively. In both experiments, the proposed method was implemented with the spatial shape transformer in Eq. (5).

### *4.1. Rigid body translation*

To verify the correctness of the proposed GC-CA framework, a rigid body translation experiment was carried out using the experimental setup shown in Fig. 3(a). The stereo measurement system was composed of two separate cameras, both of which were equipped with a monochrome imaging sensor (IDS UI-3370CP) with a resolution of 2048 × 2048 pixels and a 50 mm prime lens (Kowa Optimed). The baseline between the two cameras was about 220 mm. A flat plate with size of 100 mm × 100 mm × 6 mm was adopted as the test specimen, and a random speckle pattern was fabricated on its surface. Figure 3(b) shows a sample of the speckle pattern in the verification. The specimen was fixed on a motorized translation stage with an accuracy of 1 $\mu$m in front of the camera system, ensuring its translation can be well-controlled. The object distance from the specimen to each of the cameras was about 600 mm, and the stereo angle $\theta \approx 25$ degrees. During the experiment, the stereo image system was calibrated by the well-known Zhang's method with a planar chessboard target (11 × 8 corners, 5 mm spacing). The calibrated internal and external parameters are listed in Table 1.

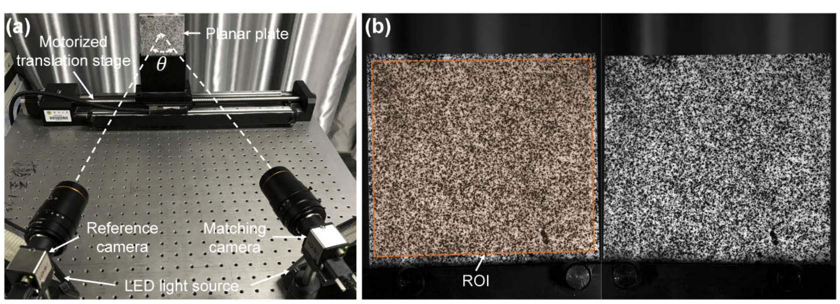


**Fig. 3.** (a) Experimental setup for rigid body translation and (b) a pair of recorded speckle images with a defined ROI.

In the experiment, the specimen was shifted by the translation stage from −1 to 1 mm with an increment step of 0.1 mm in consideration of the calibration error. At each step, a pair of speckle

Optics EXPRESS

**Table 1. Pre-calibrated internal and external parameters.**

| Parameter | Reference camera | Matching camera |
|---|---|---|
| $f_x$, $f_y$ | 10079.50, 10078.30 | 10096.50, 10101.90 |
| $c_x$, $c_y$ | 1025.37, 1015.24 | 1047.87, 1003.47 |
| $\kappa_1$, $\kappa_2$ | -0.27, 2.24 | 0.28, -2.36 |
| $\mathbf{r}$ (°) | $(-1.56,\ 24.66,\ 0.75)$ | |
| $\mathbf{t}$ (mm) | $(201.67,\ -16.67,\ 62.26)$ | |

images was recorded by the reference and matching cameras simultaneously. Subsequently, the displacements of the specimen were first calculated using the computation pipeline shown in Fig. 2, and then calculated using the stereo-DIC technique for comparison. Both techniques used the same calibration parameters, ROI and POIs, and adopted the 2D DIC method to track the POIs. The grid spacing for generating the POIs was 10 pixels and the subset size was 31×31 pixels. For 3D reconstruction in the stereo-DIC, the second-order shape function was employed to perform the stereo-matching, the linear triangulation was used to recover the object points [33]. In the GC-CA computing, a single nominal depth $\bar{d}$ = 618.52 mm was used to initialize the initial depth reconstruction. Here $\bar{d}$ was determined by the mean depth of corners on the chessboard in the first calibration pose.

Figure 4(a) shows the initial 3D shape of the specimen reconstructed by the proposed method. Intuitively, the reconstructed surface of the planar specimen looks quite fine. In order to show the quality of the reconstruction, we drew the depth distribution along the line at $Y = 0$ mm and evaluated depth variations by computing the distances between the sampled depths and the fitting line of their distribution, as shown in Fig. 4(b). The depth variation range is between -0.03 mm and 0.04 mm with a standard deviation of 0.02 mm, which seems to indicate a good planeness of the surface reconstructed by the proposed GC-CA. It is worth noting that the depth increases from left to right since the specimen was placed with its surface not vertical to the optical axis of the reference camera.

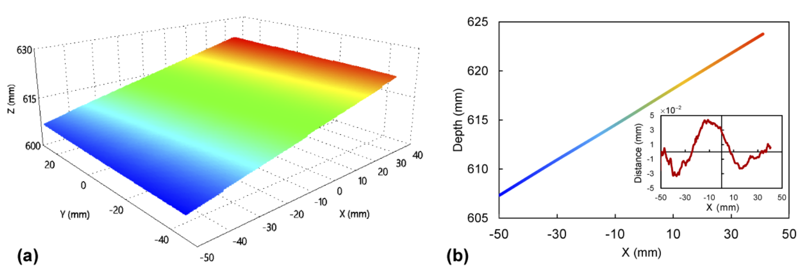


**Fig. 4.** (a) 3D shape of the planar specimen reconstructed by the proposed GC-CA framework; (b) Depth distribution along the horizontal line at $Y = 0$ mm, where the embedded curve shows the distance from the fitting line for each depth.

To show the expected performance in 3D deformation estimation, the displacements of each POI were first estimated by the proposed method and stereo-DIC, respectively. The measured translations were then obtained by averaging the total displacements of all POIs at each step for both methods and compared with the actual translations. Results are shown in Fig. 5(a). It can be seen that, although the measured translations for both methods seem in good agreement with the applied values, the results of the proposed method are closer to the actual ones than those measured by the stereo-DIC. Relative to the applied translations, the maximum absolute error and the standard deviation of the translations measured by the proposed GC-CA are 0.0095 mm

Optics EXPRESS

and 0.0030 mm. Both are lower than the corresponding errors resulting from the stereo-DIC, which are 0.0208 mm and 0.0099 mm respectively.

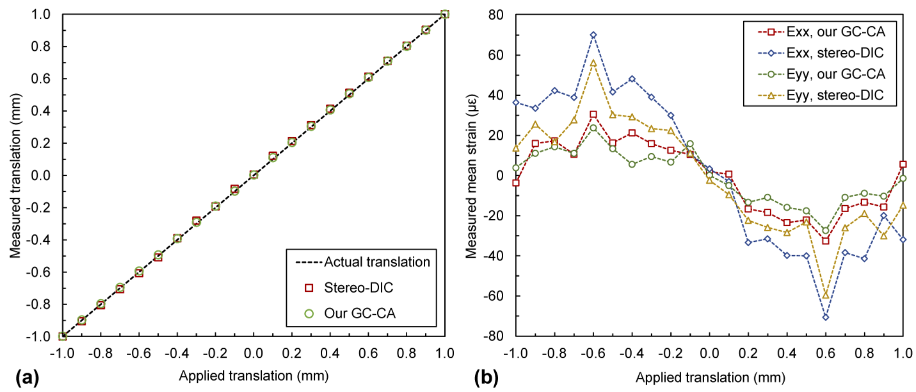


**Fig. 5.** Measured mean translations (a) and measured mean strains Exx and Eyy (b) by the proposed GC-CA and stereo-DIC, respectively.

For further investigating the performance, the strains Exx and Eyy in X and Y directions were estimated using the approach introduced in [34] with a window size of $9 \times 9$ points. The averaged strains measured by the proposed method and the stereo-DIC at each step are shown in Fig. 5(b). This figure illustrates the measured strains in both directions for the proposed method maintain a modest fluctuation with a range of about $\pm 30$ $\mu\varepsilon$. The standard deviations of the strains Exx and Eyy measured by our method are 17.3 and 12.9 $\mu\varepsilon$, respectively. By contrast, the measured strains by the stereo-DIC are clearly larger: the maximum absolute value of Exx is 70.6 $\mu\varepsilon$, with a standard deviation of 38.9 $\mu\varepsilon$; The maximum absolute value of Eyy is 59.5 $\mu\varepsilon$, with a standard deviation of 27.8 $\mu\varepsilon$. In addition, by observing the measured strains shown in Fig. 5(b), one can see that all strain distributions present similar systematic behaviors. The reason may well be both methods adopted the same calibration parameters. The comparison of the measurement errors in this experiment shows that the proposed GC-CA framework benefits to gain more accurate and precise results than the stereo-DIC method where the 3D reconstruction relies on the pre-matched geometric point correspondences. This is because the proposed GC-CA framework is capable of retrieving high-quality depth information directly from the image domain containing the underlying intensity data, so that the error propagations due to some extra operations such as explicit stereo-matching could be restrained.

### 4.2. Performance on strain measurement

To evaluate the performance of the proposed method on strain measurement, an experiment with a simply supported plate subjected to a concentrated out-of-plane load was conducted. The experimental setup is shown in Fig. 6(a). The specimen was made of Inconel 718 alloy and has a geometry in size of 150 mm $\times$ 30 mm $\times$ 3 mm. A random speckle pattern was sprayed on the surface of the specimen and recorded by a mirror-based single-camera stereo imaging system [19,24]. The focal length of the lens in the system was 12 mm, and the resolution of the camera sensor was $1280 \times 1024$ pixels. A strain gauge was attached to the middle range of the plate on the back surface. The principal direction of the strain guage was aligned along the length direction of the plate. A concentrated load was applied in the out-of-plane direction as shown in Fig. 6(a). The load was ramped up according to the readouts from the strain gauge. 12 stressed states with known strain magnitudes were achieved finally. The corresponding speckle images were recorded in phase. One sample of the captured speckle image is shown in Fig. 6(b). According to the positioning lines of the strain gauge, a virtual strainmeter was labeled with a

Optics EXPRESS

small window on the specimen surface. By doing this, the strains recorded by the strain gauge can be acted as the ground-truth for validating the proposed method.

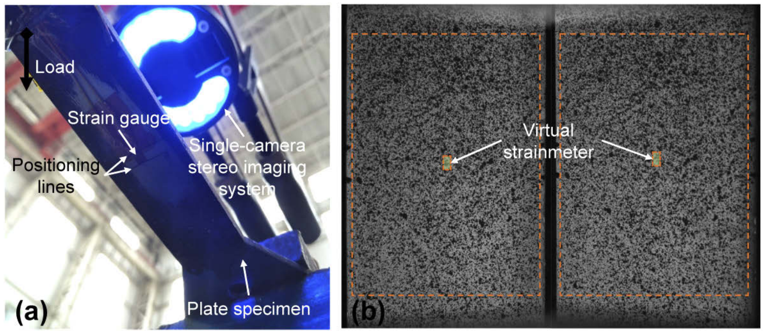


**Fig. 6.** (a) Experimental setup for 3D deformation measurement and (b) a pair of recorded speckle images with the defined virtual strainmeter.

The imaging system was calibrated by the method introduced in our previous work [25]. By averaging the calibrated internal parameters, we used the shared focal lengths $f_x$ = 2402.63, $f_y$ = 2397.86 and principle point $\mathbf{c} = (631.86, 500.92)$. The external parameters are listed as $\mathbf{r}$ = (-0.04, 48.52, -0.11) degrees, $\mathbf{t} = (77.38, 0.04, -1.88)$ mm. In the pre-calibration, 1013 pairs of features were used. After removing features close to the borders, we obtained 900 depth values in the ROI to serve as initial guesses for subsequent depth estimation. With this information, the proposed method can be used to compute the 3D displacement fields for strain analysis. The subset size in both temporal-matching and depth computing were $31 \times 31$ pixels. The average of the resulted strains within the range of the virtual strainmeter was outputted as the measured strain. For comparison, we also obtained a set of strains from the virtual strainmeter by applying the stereo-DIC technique with the same calibration parameters in the loading stage. In stereo-DIC computing, the temporal-matching method and computation parameters were the same as the proposed method, and the stereo-matching and triangulation methods were the same as the experiment in Section 4.1. The performance of the proposed method was validated by investigating the static strain errors, and comparing the measured bending strains with the ground-truth and the results of the stereo-DIC. Because of bending deformation, the strain of the upper surface was in compression while the lower surface was in tension. Therefore, in the stage of loading, we took the absolute values of the virtual strainmeter to make the signs are consistent with the ground-truth. Results are illustrated in Fig. 7.

Figure 7(a) shows the static strains evaluated by the proposed method from 49 pairs of speckle images before loading. The errors are almost less than 20.0 $\mu\varepsilon$, the maximum absolute error and the standard deviation are 20.7 $\mu\varepsilon$ and 8.7 $\mu\varepsilon$, respectively. The results suggest that the proposed GC-CA could achieve reasonable accuracy with low static strain measurement errors. Figure 7(b) shows the measured bending strain values in the deformed stages and compares against the ground-truth and the results of the stereo-DIC. It can be seen that the bending strains yielded by the proposed method are more consistent with the readouts of the strain gauge. The maximum absolute difference between the results of our method and the ground-truth is 25.6 $\mu\varepsilon$, with a standard deviation 7.3 $\mu\varepsilon$, while that for the stereo-DIC is 63.7 $\mu\varepsilon$ with a standard deviation is 19.6 $\mu\varepsilon$. All differences between the measured values of the proposed method and the gauge are less than 25 $\mu\varepsilon$ almost, presenting a consistent error level with the static errors evaluated before loading. Moreover, the strain differences seem stable and do not rise obviously with the increase of loading. By comparison, the measurement errors of the stereo-DIC, which performs 3D reconstruction with the separate stereo-matching and triangulation steps, are larger and may on an upward trend with the loading increasing. This experiment further verifies the performance

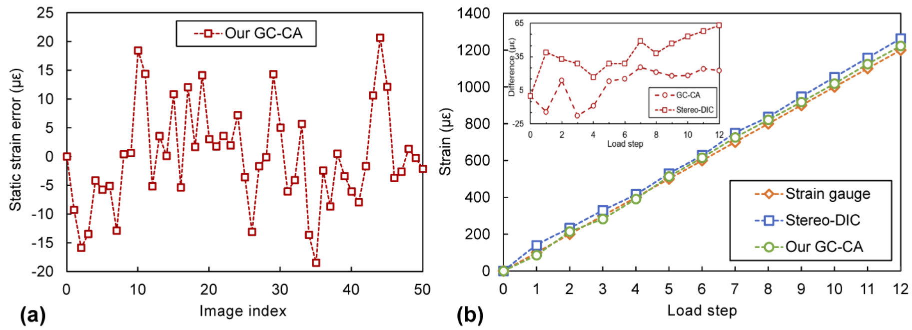


**Fig. 7.** (a) Static strain errors evaluated for the proposed GC-CA; (b) Comparison of strains measured by the proposed GC-CA with those of stereo-DIC and the ground-truth, where the difference curves show respectively the absolute errors of the GC-CA and stereo-DIC relative to the ground-truth.

of the proposed GC-CA framework, showing high-accuracy 3D deformation measurements could be expected by reconstructing the shape of deformable objects in end-to-end from the image domain.

## 5. Conclusion

We propose GC-CA, a nonlinear end-to-end depth reconstruction framework for high precision 3D deformation measurements. The proposed method links the image correlation and vision geometry to directly perform high-quality stereo depth reconstruction in the image domain, reducing the uncertainty of deformation estimation. With 3D deformation measurement experiments, we show that the feasibility of the proposed GC-CA model in 3D displacement and strain measurements, and demonstrate that it outperforms the traditional 3D reconstruction technique based on the geometric point correspondences and linear triangulation method. In addition, the efficient optimization algorithm and the implementation scheme provided in the paper ensure the proposed method can be implemented with an acceptable computational speed. This research potentially sheds new light on high-precision optical 3D shape and deformation measurements.

## Funding

National Key Research and Development Program of China (2018YFF01014200); Shanghai Post-doctoral Excellence Program (2019192); National Natural Science Foundation of China (11532005, 11672167, 11827801, 51732008, 11772092).

## Disclosures

The authors declare no conflicts of interest.

## References


1. S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," Int. J. Comput. Vis. **56**(3), 221–255 (2004).
2. Y. Su, Q. Zhang, Z. Gao, and X. Xu, "Noise-induced bias for convolution-based interpolation in digital image correlation," Opt. Express **24**(2), 1175–1195 (2016).
3. G. Bomarito, J. Hochhalter, T. Ruggles, and A. Cannon, "Increasing accuracy and precision of digital image correlation through pattern optimization," Opt. Lasers Eng. **91**, 73–85 (2017).
4. Z. Chen, X. Shao, X. Xu, and X. He, "Optimized digital speckle patterns for digital image correlation by consideration of both accuracy and efficiency," Appl. Opt. **57**(4), 884–893 (2018).
5. B. Pan, K. Li, and W. Tong, "Fast, robust and accurate digital image correlation calculation without redundant computations," Exp. Mech. **53**(7), 1277–1289 (2013).

Optics EXPRESS


6. Y. Gao, T. Cheng, Y. Su, X. Xu, Y. Zhang, and Q. Zhang, "High-efficiency and high-accuracy digital image correlation for three-dimensional measurement," Opt. Lasers Eng. **65**, 73–80 (2015).
7. X. Shao, X. Dai, and X. He, "Noise robustness and parallel computation of the inverse compositional gauss–newton algorithm in digital image correlation," Opt. Lasers Eng. **71**, 9–19 (2015).
8. J.-J. Orteu, "3-d computer vision in experimental mechanics," Opt. Lasers Eng. **47**(3-4), 282–291 (2009).
9. M. Sutton, X. Ke, S. Lessner, M. Goldbach, M. Yost, F. Zhao, and H. Schreier, "Strain field measurements on mouse carotid arteries using microscopic three-dimensional digital image correlation," J. Biomed. Mater. Res. **86A**(2), 569 (2008).
10. Z. Chen, X. Shao, X. He, J. Wu, X. Xu, and J. Zhang, "Noninvasive, three-dimensional full-field body sensor for surface deformation monitoring of human body in vivo," J. Biomed. Opt. **22**(9), 1–10 (2017).
11. Z. Hu, T. Xu, H. Luo, R. Z. Gan, and H. Lu, "Measurement of thickness and profile of a transparent material using fluorescent stereo microscopy," Opt. Express **24**(26), 29822–29829 (2016).
12. M. Malesa, K. Malowany, U. Tomczak, B. Siwek, M. Kujawinska, and A. Sieminska-Lewandowska, "Application of 3d digital image correlation in maintenance and process control in industry," Comput. Ind. **64**(9), 1301–1315 (2013).
13. J.-E. Dufour, B. Beaubier, F. Hild, and S. Roux, "Cad-based displacement measurements with stereo-dic," Exp. Mech. **55**(9), 1657–1668 (2015).
14. X. Tan, Y. Kang, and E. A. Patterson, "Experimental investigation on surface deformation of soft half plane indented by rigid wedge," Appl. Math. Mech. **37**(10), 1349–1360 (2016).
15. P. R. Gradl, "Digital image correlation techniques applied to large scale rocket engine testing," in *52nd AIAA/SAE/ASEE Joint Propulsion Conference*, (AIAA, 2016), p. 4977.
16. L. Ngeljaratan and M. A. Moustafa, "System identification of large-scale bridges using target-tracking digital image correlation," Front. Built Environ. **5**, 85 (2019).
17. F. Chen, X. Chen, X. Xie, X. Feng, and L. Yang, "Full-field 3d measurement using multi-camera digital image correlation system," Opt. Lasers Eng. **51**(9), 1044–1052 (2013).
18. D. Solav, K. M. Moerman, A. M. Jaeger, K. Genovese, and H. M. Herr, "Multidic: An open-source toolbox for multi-view 3d digital image correlation," IEEE Access **6**, 30520–30535 (2018).
19. M. E. Mohammed, X. Shao, and X. He, "Portable device for the local three-dimensional deformation measurement using a single camera," Sci. China Technol. Sci. **61**(1), 51–60 (2018).
20. T. Yuan, X. Dai, X. Shao, Z. Zu, X. Cheng, F. Yang, and X. He, "Dual-biprism-based digital image correlation for defect detection of pipelines," Opt. Eng. **58**(1), 1–13 (2019).
21. B. Chen and B. Pan, "Calibration-free single camera stereo-digital image correlation for small-scale underwater deformation measurement," Opt. Express **27**(8), 10509–10523 (2019).
22. Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Machine Intell. **22**(11), 1330–1334 (2000).
23. M. A. Sutton, J.-J. Orteu, and H. Schreier, *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts,Theory and Applications* (Springer, 2009), 1st ed.
24. X. Shao, M. M. Eisa, Z. Chen, S. Dong, and X. He, "Self-calibration single-lens 3d video extensometer for high-accuracy and real-time strain measurement," Opt. Express **24**(26), 30124–30138 (2016).
25. Z. Su, L. Lu, S. Dong, F. Yang, and X. He, "Auto-calibration and real-time external parameter correction for stereo digital image correlation," Opt. Lasers Eng. **121**, 46–53 (2019).
26. W. Tong, "Formulation of lucas-kanade digital image correlation algorithms for non-contact deformation measurements: A review," Strain **49**(4), 313–334 (2013).
27. P. Reu, "A study of the influence of calibration uncertainty on the global uncertainty for digital image correlation using a monte carlo approach," Exp. Mech. **53**(9), 1661–1680 (2013).
28. R. Balcaen, P. L. Reu, P. Lava, and D. Debruyne, "Stereo-dic uncertainty quantification based on simulated images," Exp. Mech. **57**(6), 939–951 (2017).
29. Y.-Q. Wang, M. A. Sutton, X.-D. Ke, H. W. Schreier, P. L. Reu, and T. J. Miller, "On error assessment in stereo-based deformation measurements," Exp. Mech. **51**(4), 405–422 (2011).
30. C. Zhu, S. Yu, C. Liu, P. Jiang, X. Shao, and X. He, "Error estimation of 3d reconstruction in 3d digital image correlation," Meas. Sci. Technol. **30**(2), 025204 (2019).
31. L. Yu and G. Lubineau, "Modeling of systematic errors in stereo-digital image correlation due to camera self-heating," Sci. Rep. **9**(1), 6567 (2019).
32. Z. Su, L. Lu, X. He, F. Yang, and D. Zhang, "Recursive-iterative digital image correlation based on salient features," Opt. Eng. **59**(3), 1–13 (2020).
33. R. I. Hartley and P. Sturm, "Triangulation," Comput. Vis. Image Underst. **68**(2), 146–157 (1997).
34. B. Pan, H. Xie, Z. Guo, and T. Hua, "Full-field strain measurement using a two-dimensional Savitzky-Golay digital differentiator in digital image correlation," Opt. Eng. **46**(3), 033601 (2007).