

# Evaluation of feature- and pixel-based methods for deflection measurements in temporary structure monitoring

Youyi Feng<sup>1</sup> · Fei Dai<sup>1</sup> · Hong-Hu Zhu<sup>2</sup>

Received: 18 November 2014/Revised: 29 March 2015/Accepted: 7 May 2015/Published online: 17 May 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Recent forensic studies have revealed quite a few construction site catastrophes are associated with failure of temporary structures (e.g., collapse of formworks, scaffoldings, etc.). This is particularly crucial for those located in dense urban areas where the failure of a temporary structure not only impacts the site itself but may also damage adjacent structures and cause injuries and casualties of passers-by. In this paper, the work of evaluating the optimum feature detection and matching algorithm is reported, which is the key in realizing a real-time visual sensing-based surveillance method in order to monitor the integrity and safety of on-site temporary structures. A series of experiments are designed and conducted to test three algorithms: population-based intelligent digital image correlation (DIC), David Lowe's Scale-invariant feature transform, and Hessian matrix-based speeded-up robust features. In these experiments, synthetic images through 2D geometrical translation, rotation, deformation, and illumination changes are generated to provide the sample data and ground truths. As the experiment output, the accuracy and efficiency of these algorithms are measured and compared with each other. Analyses including feasibility of interest point localization with specific regions and the error estimation with respect to real-world distances are conducted. The results show that the DIC algorithm holds the most promise to be

implemented in structural monitoring, but several challenges still remain, which call for efforts to further improvement of the technique.

**Keywords** Comparative study · Image-based methods · Construction · Safety monitoring · Remote sensing

## 1 Introduction

Temporary structures, such as formwork support, falsework, and scaffolding, are widely used in on-site production of construction projects. However, the safety of these temporary structures is not usually paid enough attention, which poses a great threat to on-site structures and workers. In fact, most of tragic failures during construction are usually the result of improperly designed, constructed, and/or maintained temporary structures. These failures will undoubtedly induce tremendous loss, delay, injuries, and casualties as the consequence.

A real-time surveillance method enabled by visual sensing may alleviate this situation. This method utilizes high-definition cameras to monitor the integrity and safety of a site temporary structure while it is undergoing excitements during site ongoing operations. This method has the advantage of being cost-effective and easy-to-deploy. In essence, this study intends to develop a cost-effective yet applicable solution to measure deflections of temporary structures. The failure of structures can be associated with large deflection or deformation on the surface of the structures. Promptly obtaining the deflection or deformation may help assess whether a structure is under threat. For example, if the observed deflection is approaching or exceeding certain threshold that the structure can withstand, it may indicate the structure is under high risk of collapse.

---

✉ Fei Dai  
fei.dai@mail.wvu.edu

<sup>1</sup> Department of Civil and Environmental Engineering, West Virginia University, P.O. Box 6103, Morgantown, WV 26506, USA

<sup>2</sup> School of Earth Sciences and Engineering, Nanjing University, Nanjing 210023, China

## 2 Background

### 2.1 Temporary structure failures

It is crucial to guarantee the safety and integrity of temporary structures during the process of construction. Otherwise, failure of temporary structures may occur and lead to severe consequences to the project. Table 1 provides six examples with the injuries and fatalities due to temporary structure failures in recent years.

Recent forensic studies have revealed quite a few construction site catastrophes are associated with failure of temporary structures (e.g., collapse of formworks, scaffoldings, etc.) [1]. It is particularly crucial for those that are located in dense urban areas where the failure of a temporary structure not only impacts the site itself but may also damage adjacent structures and cause injuries and casualties of passers-by. In addition, some geological disasters, such as the collapse during the progress of deep excavation and tunneling, will also result in tragic calamities [2]. As an inevitable consequence, the collapse of the temporary structures will result in enormous loss, injuries, and casualties.

To reduce the painful casualties induced by temporary structure collapses, the current methods to ensure the safety of temporary structures include regular manually inspection [3], measuring instrument-based surveying, such as sensor networks, total station, laser scanner [4–6], and visual sensing techniques, such as digital image correlation (DIC), interest feature detecting, extraction, and tracking [7].

### 2.2 Review of current methods and limitations

#### 2.2.1 Regular manual inspection

Normally, site engineers periodically check the compliance of structures manually with specifications and rules regulated by governments or agencies. The engineers will conduct a visual inspection to check the apparent deficiencies and damage to determine if the temporary structure meets the specific safety criteria. However, there

exists a major defect with this practice due to its incapability of continuous monitoring of on-site structures. Engineers usually perform the inspection work before and after normal construction hours. This means that the safety of temporary structures is not secured when they are in service [8], leading to the safety of construction workers still under potential threat during construction operations.

#### 2.2.2 Measuring instrument-based surveying

In recent years, the most commonly used instruments for structure monitoring are wire/wireless sensors [9, 10]. The operators install the wire/wireless sensors onto the structures to be monitored, whereby they can collect the information regarding the target position changes. This method can achieve sufficient accuracy performance [11]. However, despite the fact that the development of wireless sensing technique may reduce the extra cost of the wire transmission needed to some extent, the convenience of switching operation between different monitoring targets is still an issue. Besides, when facing relative large-scale applications, the number of sensors needed for installation and uninstallation will be another issue which calls for extra efforts of the users. In addition, since the sensor networks method can only detect the deflections and strains of those positions where the sensors were put onto, the full-field deflections and strains of the integral target still cannot be achieved.

Other existing methods may overcome the mentioned deficiency of sensor networks, one of which is laser scanning. Laser scanning generates 3D point cloud of the monitored target. Based on the point cloud before and after deformation, the spatial deflection information of the target will be obtained [12]. The main drawback of this method is that the expense of laser scanner used for surveying is normally over thousands of dollars, which causes this method actually not practical for research with relatively low cost to monitor the temporary structures on a construction site.

Another measuring instrument is total station, which can also be applied to monitor the safety of the

**Table 1** Examples of temporary structure failures

Time	Location	Injuries/fatalities	Accident depiction
3/09/02	Chicago, IL, US	8 injuries, 4 killed	John Hancock Center, suspended scaffold collapse
2/22/07	Beijing, China	16 injuries, 3 killed	Formwork collapse during concrete pouring
11/19/08	Oklahoma City, OK, US	10 injuries	Two of the temporary stages collapsed
5/20/09	Middletown, DE, US	2 injured, 1 killed	One side of the trench collapsed suddenly
1/12/10	Hefei, Anhui, China	7 injuries, 8 killed	Formwork support collapse
5/8/11	Toronto, Ontario, Canada	7 injuries, 1 killed	The main stage of Big Valley Jamboree collapsed

temporary structures. In this method, the special markers are placed on the target, and then the total station will be operated to record the spatial coordinates of the markers to acquire the target's position changes information [13]. However, using total station faces the same problem as applying sensor networks. That is, it cannot achieve full-field measurement of deflections and strains.

## 2.3 Visual sensing techniques

### 2.3.1 Overview

Considering the above-mentioned unfavorable factors of the conventional methods, visual sensing techniques are employed to addressing issues that exist in temporary structure monitoring. Visual sensing techniques, namely, applying visual sensors (such as digital video cameras, industrial cameras, etc.) to capture and record the spatial position information of a target (normally, the output format will be digital image or video), and then based on relevant visual principles and algorithms to recover the real-world spatial information of the target. After obtaining the necessary spatial position information, the deflection and strain of the target can be computed as the output [14]. Specifically, the camera station is fixed when monitoring a target. So any position changes of the target will lead to the position changes of features in consecutive frames. The deflection of the target therefore can be calculated by tracking the position of the features in images over time before and after deflection. After matching the correspondences of the same features in the images before and after deflection, the distance in pixel each feature deviated composes the in-plane deflection. This method has several advantages when compared to the conventional methods. Firstly, it can generate the full-field deflection and strain of the target. Secondly, it is convenient to operate this technique periodically and apply to different scenarios [15]. As a result, this visual sensing-based method can be a good option for use in temporary structure monitoring.

The basic principle of this method lies in detecting and tracking of interest features in the sequencing image frames to capture the target point coordinates' changing values. Then, based on the consistent coordinates of the points in each frame, the monitored target's deflection will be computed. In specific, the accuracy of the monitoring method is entirely associated with the interest points' location continuously varying in the frames. Hence, the key task to achieve a high-accuracy monitoring method is to obtain the most appropriate feature detecting and tracking algorithm to acquire the feature points' coordinates as accurately as possible [16].

Currently, a series of algorithms have been developed to detect and track feature points along image streams. These algorithms can be categorized into two types [17].

The first type is associated with feature-based pixel-level matching. In this type, Scale-Invariant Feature Transform (SIFT) proposed by David Lowe is known as the most typical algorithm [18]. The features are scale-invariant to image scaling, rotation, and affine transformation. This favorable property makes this algorithm one of the most useful algorithms for feature point detection and tracking. Speeded-up robust features (SURF) is another feature-based matching algorithm proposed by Bay et al. [19]. It inherits the scale-invariant feature, and its running efficiency has been proved to be much higher than SIFT [20]. The reason that we select these two algorithms in our study is because previous research has revealed that the SIFT and SURF detectors and descriptors outperform other detectors and descriptors such as histogram of oriented gradients (HOG) and gradient localization oriented histogram (GLOH) [21]. Also, it has been reported that SIFT and SURF are good candidates for taking geometric measurements in civil infrastructure surveying-related applications [22].

The second type is related to pixel-based sub-pixel-level matching. Digital image correlation belongs to this type, which has been applied in applications such as mechanical deformation detection. Research has proved this algorithm has great potential to detect deflection and strains in mechanical field [23]. However, the performance of this algorithm in civil engineering has still not been revealed. There are distinct differences for this technique to be used in civil engineering and mechanical settings. For example, speckle patterns are usually applied for DIC in experimental testing of mechanical applications [24], whereas the speckle patterns are actually not feasible to be utilized in some civil engineering applications. As a result, to further identify the potential of the DIC algorithm for use in measuring deflections of civil structures, further studies are needed to compare the DIC algorithm with the feature-based matching algorithms of SIFT and SURF.

The following sections will present the principles of the three algorithms of SIFT, SURF, and DIC. This will be followed by the statement of the research problems and objective. Then, detailed experimental design, implementation, results, and analysis will be delineated.

### 2.3.2 Scale-invariant feature transform (SIFT)

To achieve feature point detection and matching, there are four main procedures involved in the SIFT algorithm.

- Extreme point (local maxima or minima) detection of the difference of Gaussians (DoG), applying Gaussian

differential function to extract scale-invariant interest points;

- *Key point localization* At the location of interest points, each potential key point is checked. Those points that have low contrast values or are poorly localized along edges are removed.
- *Orientation for key points* The orientations of the key points are calculated and dominant orientations are identified and assigned. If a key point has multiple dominant orientations, an additional key point is created at the same location and scale for each additional dominant orientation.
- *Key point description* A SIFT description is computed for each key point to make the key point distinctive.

The feature points detected and described by the SIFT detector and descriptor have the following properties [25]. First, the detected features are invariant to scale and rotation, since the key points are selected as local maxima or minima across scales and related to their dominant orientations. Also, the detected features are robust to changes in illumination. In addition, the features are highly distinctive and easy to extract. The favorable properties make this algorithm one of the most useful algorithms for feature point detection and matching in a number of related areas [26].

### 2.3.3 Speeded-up robust features (SURF)

The SURF algorithm, which was first presented by Bay et al. [19], adopted the concept of the Hessian matrix and approximated the determinant of the Hessian matrix with two box filters. Based on the image response values to the filters, the key points can be localized with a non-maximum suppression schema by interpolating the maxima of the determinant of the Hessian matrix across scales.

After the key point localization, based on the distributions of the intensity content within the local regions of the points, the descriptions of the key points can be generated. Typically, in SURF algorithm, the point description is designed as a vector with 64 elements. To build up such a SURF description, firstly the dominant orientation of each key point is identified based on the wavelet transform [27]. Once the dominant orientation is determined, a square window will put at the point. Then rotating the window is performed so that it is oriented along the point's dominant orientation. Afterwards the window is divided into 16 small square windows. Then, the horizontal and vertical response values on account of Haar-wavelet transform in each small window will be summed up. The sums of the response values in two directions for all 16 small windows putting together generate the description vectors used to characterize the key points.

SURF algorithm has several distinctive properties. Firstly, the algorithm relies on integral images to reduce the computational time, which makes its running efficiency higher than SIFT [28]. In addition, it is robust against image scale and rotation. However, the performance of SURF algorithm is not as good as that of SIFT algorithm on the invariance to image illumination and viewpoint changes. So far, the effectiveness of SURF algorithm has been validated in several related tasks, such as object detection, recognition and tracking [29], and 3D reconstruction [30].

### 2.3.4 Digital image correlation (DIC)

Digital image correlation is known as a non-contact full-field 2D/3D measurement method, which computes the changes (displacement, deformation, and strain) in images. It employs image registration and tracking techniques to measure the planar or spatial deflection and deformation within a series of continuous image frames [31].

The DIC algorithm implemented and tested in our research applies the zero-mean normalized cross-correlation (ZNCC) criteria, which is insensitive to scale and image illumination changes [32]. The ZNCC criteria are described as bellow.

$$C(p') = 1 - \frac{\sum_{x=-M}^M \sum_{y=-M}^M [f(x, y) - f_m][g(x', y') - g_m]}{\sqrt{\sum_{x=-M}^M \sum_{y=-M}^M [f(x, y) - f_m]^2} \sqrt{\sum_{x=-M}^M \sum_{y=-M}^M [g(x', y') - g_m]^2}},$$

where the  $f(x, y)$  and  $g(x', y')$  are the corresponding gray values of the deformed reference subsets,  $x, y$  are the coordinates anchored at the center of the reference subset coordinate system, and

$$f_m = \frac{1}{(2M + 1)^2 \sum_{x=-M}^M \sum_{y=-M}^M f(x, y)}$$

$$g_m = \frac{1}{(2M + 1)^2 \sum_{x=-M}^M \sum_{y=-M}^M g(x', y')}.$$

Then, the values of  $f_m$  and  $g_m$  are calculated as the average gray values of points in the two subsets, and  $p'$  is described as the deformation vector, which reveals the relationships between coordinates  $(x, y)$  and  $(x', y')$ . In addition, the point with the coordinates  $(x, y)$  in the reference subset after deformation can be represented by the first-order shape function:

$$x' = x + u + \frac{\partial u}{\partial x}x + \frac{\partial u}{\partial y}y,$$

$$y' = y + v + \frac{\partial v}{\partial x}x + \frac{\partial v}{\partial y}y,$$

where  $u$  and  $v$  are the displacement components of reference subset center on  $x$  and  $y$  directions. The expressions

$\partial u/\partial x$ ,  $\partial u/\partial y$ ,  $\partial v/\partial x$  and  $\partial v/\partial y$  are the displacement gradient components and  $p' = [u, v, \partial u/\partial x, \partial u/\partial y, \partial v/\partial x]$  is calculated as the corresponding parameter vector.

The first-order shape function proposed above can deal with translation, rotation, shear, strains, and their combinations of the subsets, and it can recover all necessary deflection and deformation information for measurement in our research.

## 2.4 Problem statement and objective

Our overarching research goal is to establish a visual sensing-based monitoring method that can be real-time and continuously monitor the full-field displacements and strains of load-bearing members of site structures. The method is enabled by video/industrial cameras to facilitate the monitoring in the entire construction process to avoid accidents to the maximum extent. This will also compensate for the deficiency of design and construction of temporary structures. It entails utilizing digital cameras to capture a series of target image frames and applying image processing algorithms to compute the target's deformation so as to alert the site engineers when the deformation is in large scale and may cause an accident.

However, the accuracy of such image-based monitoring method is associated with identification of the interest points' location in images. Therefore, the key task to achieve a high-accuracy monitoring method is to get the optimal feature matching algorithm that is capable of detecting position changes of feature points over images as accurately as possible.

The feature-based and pixel-based methods as mentioned in previous section are commonly used in detecting and matching variance that occurs on structure members. However, which one is the most appropriate for implementation of temporary structural monitoring is unknown in terms of accuracy and efficiency. To this end, this paper makes an attempt to compare the accuracy and efficiency of these algorithms.

## 3 Experiment

### 3.1 Experiment design

To test the performance of the three algorithms, experiments are designed and carried out. 2D images are selected and used to generate synthetic 2D image datasets for experiment. The synthetic images are used to provide the ground truths, which are generated by the implemented MATLAB geometric transformation and deformation

functions based on the original images. In this paper, the in-plane deflection distances of the corresponding interest feature points from one synthetic image pair are the criteria to reflect the accuracy performance of these algorithms. The experiments are designed to test their performance in dealing with feature point detection and matching under different scenarios. The performance criteria include feature detection accuracy, algorithm operation efficiency, and feasibility. The geometric transformations include translation, rotation, illumination, and sinusoidal function deformation. Three algorithms are tested. The experimental results are the deflections (distances between the interest feature points in the original image and that of the transformed synthetic image). The deflection values are then compared with their corresponding ground truths to analyze the detection performance.

### 3.2 Dataset

The dataset applied in the experiments includes three different scenarios which are shown in Fig. 1. The scene of image (a) is a part of a bridge in the field; the scene of image (b) is a constructed mock-up bridge in the laboratory; the scene of image (c) is a building structure at a construction site.

#### 3.2.1 In-plane translation

This experimental group is designed to test the performance of the algorithms to deal with the in-plane translation (displacement) of the three scenarios. The deflection direction of the synthetic images is vertical, and the values of deflection are 3, 5, and 8 pixels, respectively, as shown in Fig. 2.

#### 3.2.2 In-plane rotation

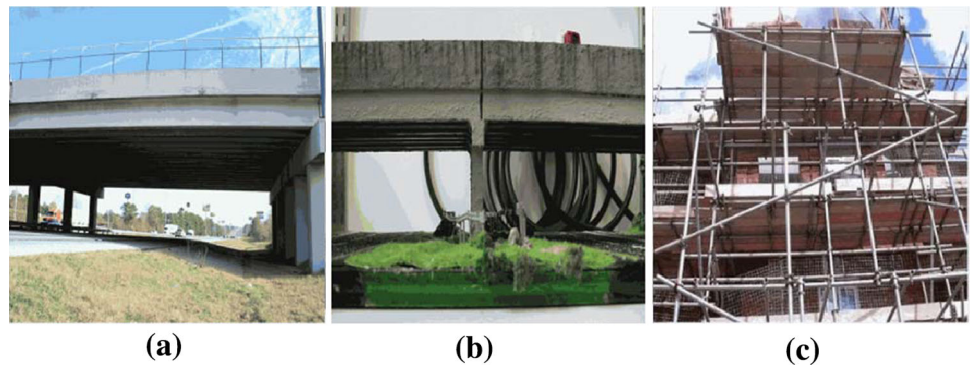
The dataset in this group is designed to test the performance of the algorithms in dealing with the in-plane rotation of the scenarios. The rotate direction of the synthetic images is clockwise, and the rotation angles are 0.5, 1, and 1.5°, respectively, as shown in Fig. 3.

#### 3.2.3 Illumination changes

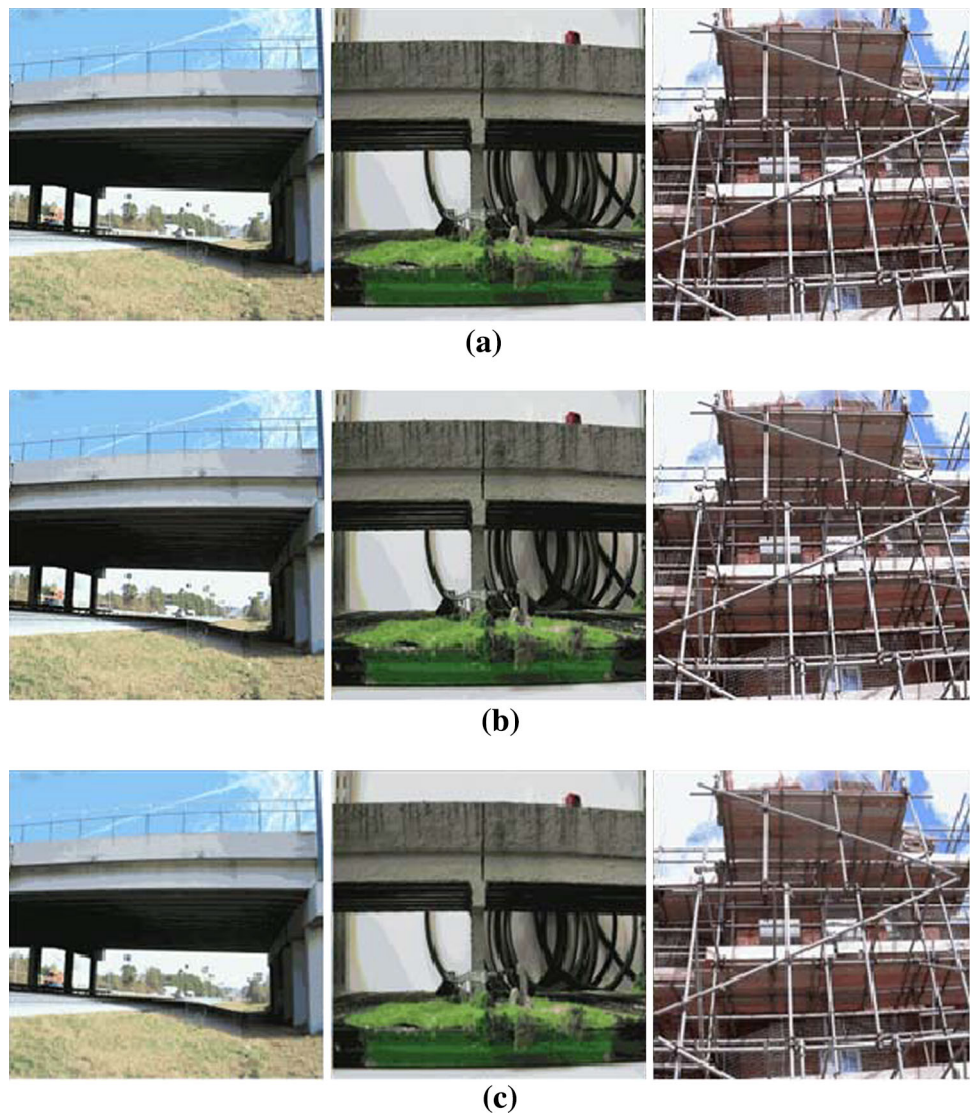
The experiment is designed to assess the algorithm performance in dealing with illumination changes when the pictures are captured under different lighting conditions. This is essential since the algorithms are expected to have the ability to monitor the scenes under different illuminations during daytime. This group of experiment is conducted by utilizing the synthetic images with 5 pixels



**Fig. 1** Dataset of original images: **a** a bridge, **b** a mock-up bridge, and **c** a building under construction



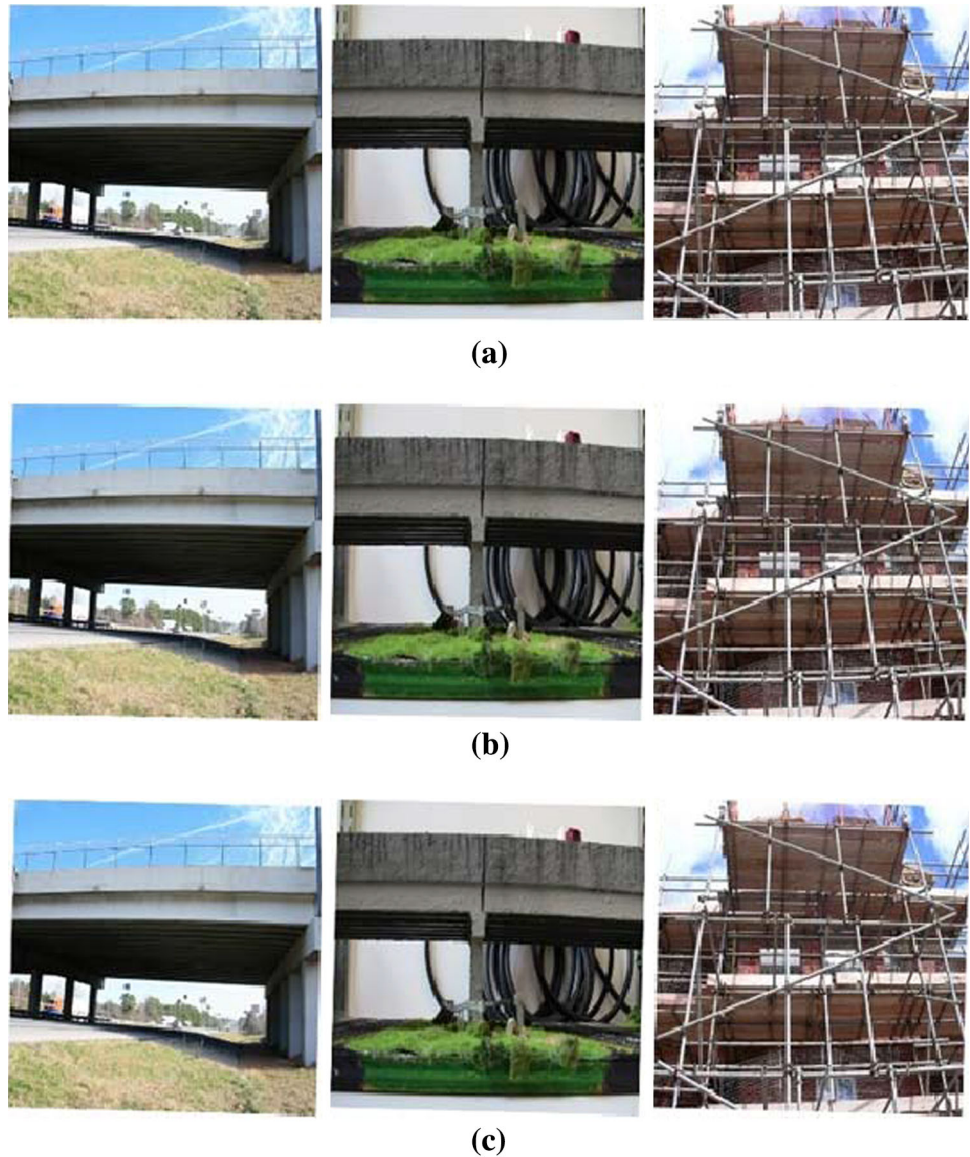
**Fig. 2** **a** 3 pixels deflection scenarios; **b** 5 pixels deflection scenarios; **c** 8 pixels deflection scenarios



deflection, which are processed in the Xnview software to reduce and increase the lightness of different images. The less illumination group is generated by reducing 50 lux

from the original image group. In contrast, the more illumination experimental group is increased 50 lux based on the original image group, as shown in Fig. 4.

**Fig. 3** a 0.5° rotation scenario; b 1° rotation scenario; c 1.5° rotation scenario



3.2.4 In-plane deformation

The experimental groups presented above are all about linear transformations of the scenarios. However, as well known, in the real-world, the observed targets cannot always be considered to perform as linear transformation. Therefore, this group is designed specifically to address the non-linear transformation scenarios, also known as deformation, which is a widespread issue in construction temporary structures.

In this experimental group, the deformed images are generated by the specified deformation functions:

$$\begin{cases} u = x \\ v = y + [\mu \times \sin(2 \times \pi \times x)/(h - 1)] \end{cases} \quad (1)$$

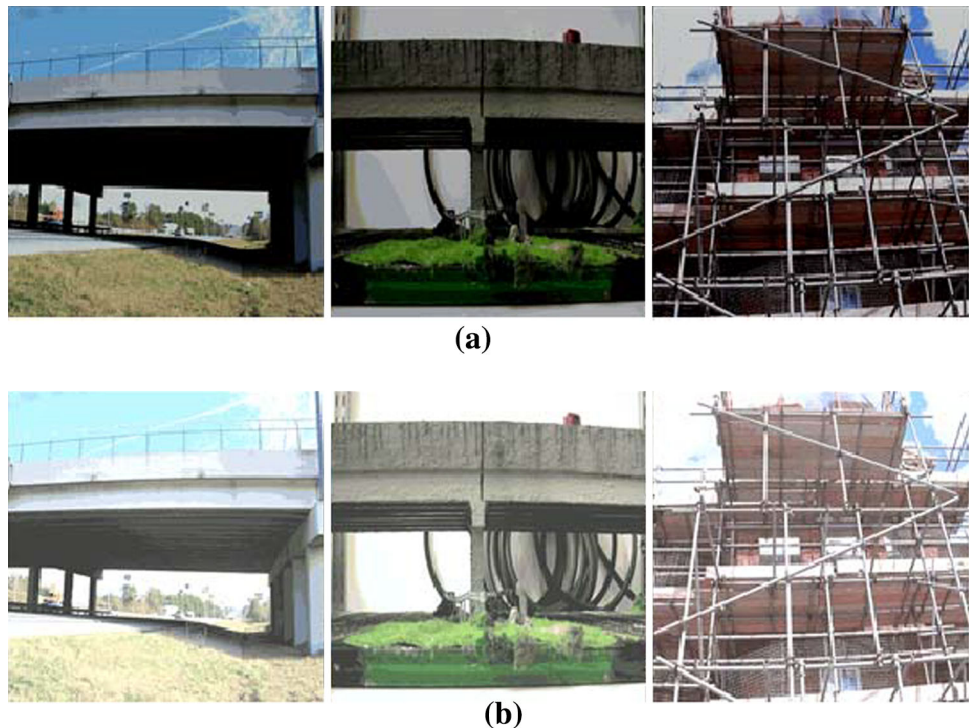
$$\begin{cases} u' = x + [\mu \times \sin(2 \times \pi \times y)/(w - 1)] \\ v' = y \end{cases} \quad (2)$$

The functions above are used to perform sinusoidal deformation to the target scenarios. Equation (1) is used to add vertical sinusoidal wave to the target scenarios. Similarly, Eq. (2) adds horizontal sinusoidal wave to the target scenarios.

In the functions, the  $u$ ,  $v$ ,  $u'$ , and  $v'$  are the coordinates of the deformed image points, and  $x$  and  $y$  are the coordinates of the original image points. The deformation scale factor  $\mu$  is set to 5.0 in the experiments.  $\pi$  is the circumference ratio, and  $h$  and  $w$  are the height and width of the images.



**Fig. 4** **a** 50 Lex less illumination images; **b** 50 Lex more illumination images



## 4 Experimental results

### 4.1 Algorithm accuracy comparison

This section presents the statistical comparison results that are generated by processing the images of different scenarios with the three algorithms. The data in Table 2 are overall measurement error percentages of the different scenarios.

#### 4.1.1 Translation

Figure 5 shows the measurement errors of the three algorithms for different scenarios and different translation values. From Fig. 5, we can see that the SURF algorithm has the largest measurement error percentages among these three algorithms. SIFT algorithm performs much better than SURF, but worse than DIC. As a result, regarding translated scenarios, DIC has the best measurement performance among the three algorithms.

#### 4.1.2 Rotation

The measuring error percentages of the three algorithms for rotation scenarios are presented in Fig. 6. From the plots in Fig. 6, the SIFT and SURF algorithms have comparable measuring performance, whereas DIC algorithm achieved much better accuracy than SIFT and SURF when processing the rotation scenarios. However, Fig. 6 also shows

that the measuring error percentages of the three algorithms for rotation group are all greater than 15 %, which indicates that all three algorithms have some difficulties in dealing with rotation scenarios when compared with the translation group.

#### 4.1.3 Changing illumination

Figure 7 presents the measurement results of the three algorithms when changing the light condition of the scenarios. Figure 7 shows that SURF algorithm is much more sensitive to light condition compared with SIFT and DIC algorithms. Its average error percentage is around 5 %, which is much larger than that of SIFT and DIC algorithms. Besides, from Fig. 7, the accuracy performance of SURF and SIFT algorithms varies a lot when processing different scenarios with the same light condition. In contrast, DIC achieved much better results.

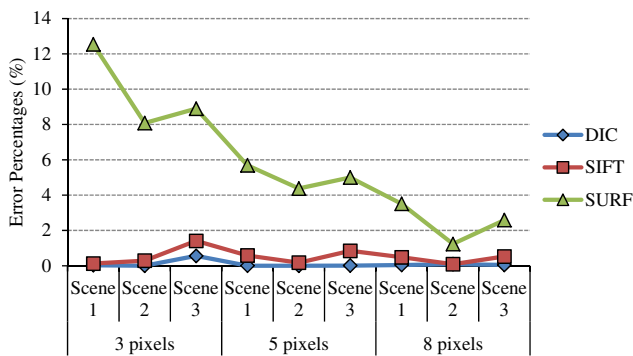
#### 4.1.4 Deformation

The results of the three different algorithms for deformation scenarios are presented in Fig. 8, from which, it can be observed that DIC achieved the best performance among the three algorithms; its error percentages lie between the intervals of 10 % and 15 %. SURF algorithm performed better than SIFT algorithm. However, all the three algorithms still need further improvement to gain better measuring accuracy for deformation scenarios.

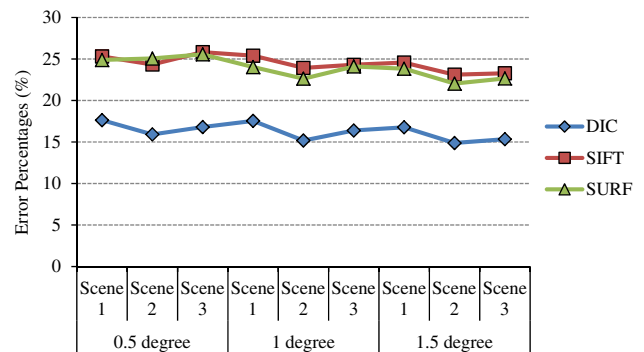


**Table 2** Accuracy comparison result

Translation (%)	3 pixels			5 pixels			8 pixels		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	0.0311	0.0003	0.5587	0.0045	0.0001	0.0124	0.0443	0.0883	0.0464
SIFT	0.1252	0.2914	1.4079	0.5825	0.1743	0.8446	0.4822	0.0890	0.5279
SURF	12.5417	8.0829	8.8976	5.6888	4.3751	5.0000	3.5072	1.2331	2.5856
Rotation (%)	0.5°			1°			1.5°		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	17.6293	15.9105	16.7992	17.5450	15.1788	16.3858	16.7685	14.8734	15.3340
SIFT	25.3056	24.3408	25.8251	25.3970	23.9278	24.3075	24.5719	23.1111	23.2874
SURF	24.8862	25.0558	25.5584	24.0299	22.6295	24.1035	23.8335	22.0336	22.6595
Illumination (%)	Less illumination			Normal illumination			More illumination		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	0.0043	0.0813	0.0023	0.0045	0.0001	0.0124	0.0045	0.0001	0.0140
SIFT	0.5825	0.1803	0.8446	0.5825	0.1744	0.8446	0.5825	0.1803	0.8446
SURF	6.1731	4.3751	5.6578	5.6888	4.3751	5.0000	5.7403	4.3751	5.2190
Deformation (%)	Scene 1			Scene 2			Scene 3		
	DIC			SIFT			SURF		
DIC	17.5012			18.839			18.5891		
SIFT	19.9427			26.1856			27.329		
SURF	19.5472			25.6559			26.3965		



**Fig. 5** Measurement error percentages of the three algorithms for translation scenarios

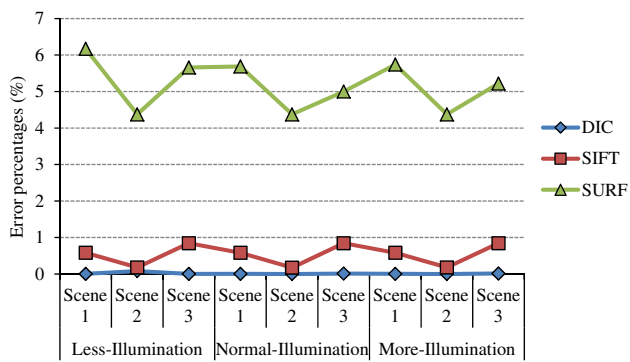


**Fig. 6** Measurement error percentages of the three algorithms for rotation scenarios

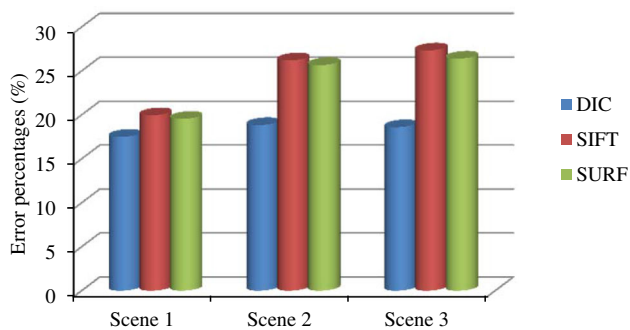
### 4.2 Algorithm efficiency comparison

After the accuracy comparison, the algorithms' running times were compared to evaluate their efficiency performance. The running times for the algorithms in processing different testing groups are recorded in Table 3.

Figure 9 shows the running times of the three algorithms for different scenarios. From Fig. 9, for translation and illumination groups, SURF and SIFT algorithms have the best efficiency performance. The running time of SURF is within 1 s, and for SIFT, it is around 1 s. Both SURF and SIFT can be considered as near real time for processing. However, for DIC algorithm, the running time is much



**Fig. 7** Measurement error percentages of the three algorithms for changing illumination scenarios



**Fig. 8** Measurement error percentages of the three algorithms for deformation scenarios

larger than that of SURF and SIFT. In particular, for the rotation scenarios, the running time of DIC is more than 10 s for each scenario.

### 4.3 Error estimation in real-world scene

To shed light on the performance of the algorithms on metric measurements, we further calculate the pixel errors into real-world values given the camera capture distance and its focal length. Based on imaging principle of digital camera, the measuring pixel-based error can be converted into real-world Euclidean metric error by given different image resolutions. Firstly, the average pixel-based errors for each algorithm with different types of transformations were computed as shown in Table 4.

Then, the pixel error can be calculated into real world with given capturing distance 5 m for the camera. The camera we used in our experiment is Canon 5D mark III which is equipped with a  $36 \times 24$  mm full frame CMOS. The lens used for the camera in the experiment has a 30 mm fixed focus. Based on different image resolutions, the real-world errors calculated are shown in Table 5. The unit of the errors presented in the table is in millimeter,

which is calculated by specifying the camera capture distance and given different image resolutions.

The experimental results using different image resolutions for the three algorithms are presented in Fig. 10. From Fig. 10, when comparing the real-world accuracy between different algorithms, we can see that DIC achieved the best accuracy performance among the three algorithms. SURF presented the worst accuracy performance. Especially, when the image resolution becomes  $720 \times 480$  pixels, the real-world error of SURF dealing with translation scenarios is over 2 mm that is much higher than the errors of DIC and SIFT. In particular, all of the three algorithms seem to have difficulties in processing the rotation and deformation groups when the image has a relatively small resolution.

## 5 Discussion

### 5.1 Algorithm accuracy performance

According to Table 2, we can observe that DIC can achieve the best accuracy performance among all of the testing groups including translation, rotation, illumination changes, and deformation. SIFT led to more accurate measurements than SURF in scenarios of translation and illumination changes. In terms of the rotation and deformation, the performance of SIFT is close to the performance of SURF. This indicates that they can achieve the similar accuracy when dealing with different rotation and deformation scenarios.

### 5.2 Running efficiency

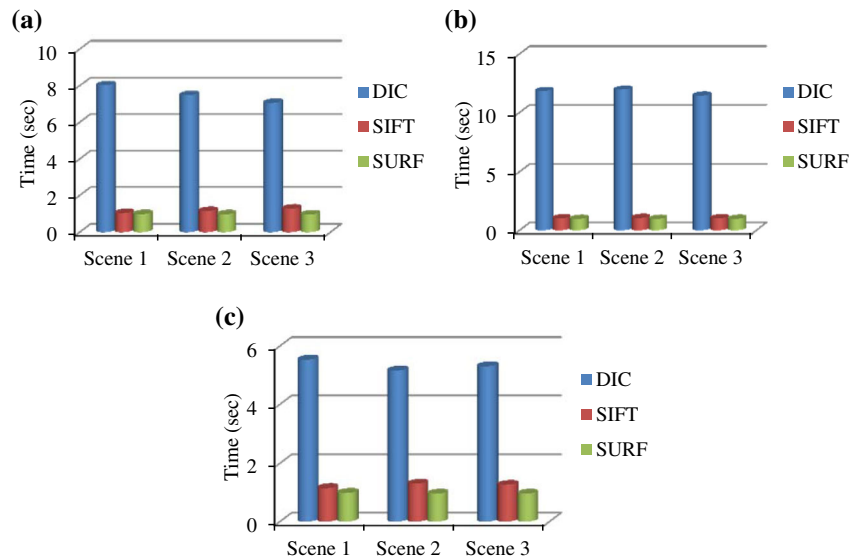
Table 3 presents the running time of the algorithms in processing different image scenes. All the images processed in the efficiency experiments have the resolution of  $500 \times 500$  pixels. As we see in Table 3, each algorithm's running time for different scenes (translation, rotation, and illumination change) has little difference. This indicates that the scene difference has relatively small influence on the operation efficiency of the three algorithms.

However, the running time of different algorithms in processing the same scenarios is quite different. As shown in Table 3, SURF achieved the best time efficiency among three algorithms. SIFT is the second. The DIC is the worst. The running time of SURF algorithm is around 0.2 s and the running time of SIFT algorithm is around 1 s. These can be considered to be real time or near real time for image data processing. However, the running time of DIC varies with different scenarios, around 5 s in illumination change and translation groups, and 12 s in rotation groups.

**Table 3** Efficiency comparison result

Translation (s)	3 pixels			5 pixels			8 pixels		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	5.979	4.894	4.148	5.733	5.285	4.584	12.356	12.253	12.435
SIFT	1.046	1.05	1.023	1.045	1.342	1.277	1.025	1.062	1.543
SURF	0.982	0.981	0.961	0.983	0.958	0.957	0.982	0.975	0.960
Rotation (s)	0.5°			1°			1.5°		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	11.446	11.666	9.982	11.687	11.788	11.927	12.396	12.408	12.396
SIFT	0.99	0.98	0.991	1.035	1.083	1.053	1.042	1.098	1.022
SURF	0.941	0.959	0.944	0.982	0.955	0.997	0.995	0.997	0.994
Illumination (s)	Less illumination			Normal illumination			More illumination		
	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3	Scene 1	Scene 2	Scene 3
DIC	5.733	5.285	4.584	4.974	5.261	5.986	5.898	4.962	5.365
SIFT	1.045	1.342	1.277	1.126	1.235	1.338	1.261	1.325	1.167
SURF	0.983	0.952	0.958	0.978	0.951	0.952	0.979	0.956	0.951

**Fig. 9** Efficiency comparisons of the three algorithms for **a** translation scenarios, **b** rotation scenarios, and **c** illumination change scenarios



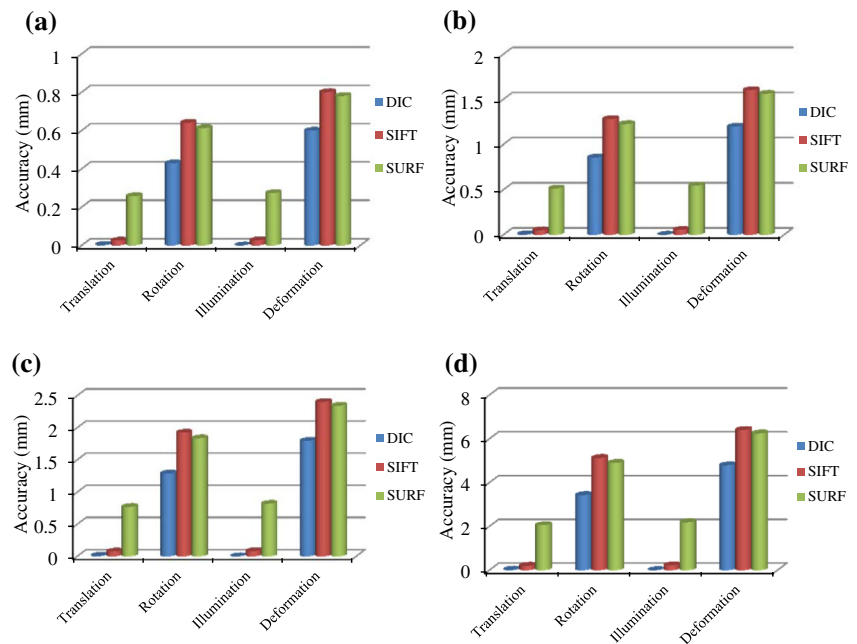
**Table 4** Measurement average pixel-based error of different algorithms

Average error (pixel)	Translation			Rotation			Illumination		Deformation
	3 pixels	5 pixels	8 pixels	0.5°	1°	1.5°	Less	More	Sinusoidal
DIC	0.0059	0.0003	0.0048	0.2871	0.3652	0.5841	0.0015	0.0003	0.5752
SIFT	0.01825	0.0267	0.0293	0.4260	0.5447	0.8763	0.0268	0.0268	0.7692
SURF	0.29522	0.2511	0.1954	0.4014	0.5560	0.8074	0.2701	0.2556	0.7498



**Table 5** Real word errors with 5 meters capture distance

Real-world error estimation (mm)	Translation	Rotation	Illumination	Deformation
Resolution: 5760 × 3840				
DIC	0.0038	0.4293	0.0009	0.5992
SIFT	0.0258	0.6413	0.0279	0.8013
SURF	0.2575	0.6128	0.2738	0.7810
Resolution: 2880 × 1920				
DIC	0.0076	0.8586	0.0019	1.1983
SIFT	0.0516	1.2826	0.0558	1.6025
SURF	0.5150	1.2256	0.5476	1.5620
Resolution: 1920 × 1280				
DIC	0.0114	1.2879	0.0028	1.7975
SIFT	0.0773	1.9240	0.0837	2.4038
SURF	0.7726	1.8383	0.8214	2.3430
Resolution: 720 × 480				
DIC	0.0304	3.4344	0.0074	4.7934
SIFT	0.2062	5.1306	0.2233	6.4102
SURF	2.0601	4.9022	2.1903	6.2481

**Fig. 10** Real-world scenario testing results of the three algorithms using image resolutions of **a** 5760 × 3840 pixels, **b** 2880 × 1920 pixels, **c** 1920 × 1280 pixels, and **d** 720 × 480 pixels

### 5.3 Metric measurement error estimation

In Table 5, the image pixel errors were computed into real-world metric measurements. As to translation group, when the camera capture distance is 5 m from the scene, the pixel errors recalculated into real world are approximately 0.004 mm for DIC with the image resolution of 5760 × 3840 pixels. As the image resolution reduces, the error increases to 0.03 mm with the 720 × 480 pixel resolution. Similar results were obtained in the rotation, illumination, and deformation groups. As a result,

regarding the same algorithm, the real-world accuracy has positive correlation relationship with the image resolution.

In addition, the performance of each algorithm in processing different transformation scenarios (i.e., translation, rotation, illumination and deformation) can be compared. For the translation groups, DIC obtained the best in accuracy. The real-world errors of DIC are all within 0.1 mm. SIFT can result in acceptable accuracy that is within 1 mm for all image resolutions. SURF has the worst performance in dealing with the same translation group. Similar observations happened to the illumination groups. As to the

rotation and deformation groups, the accuracy performance of all three algorithms is not as high as translation and illumination groups. However, among these three algorithms, DIC still results in better accuracy than SIFT and SURF. In particular, all of the three algorithms seem to have issues in processing the rotation and deformation groups when the image has a relatively small resolution. For example, in the deformation group, when the image resolution is  $720 \times 480$  pixels, the accuracy performance of all three algorithms is over 4 mm. This being observed, the algorithms call for further improvement to address this problem.

## 6 Conclusions

This paper explored the optimal detection and tracking algorithm for establishing a temporary structure monitoring method. A series of experiments were conducted to test three algorithms: DIC, SIFT and SURF. In the experiments, synthetic images, including 2D geometrical translation, rotation, deformation, and illumination changes, were generated to provide the sample data and ground truths. As the experimental output, the accuracy and efficiency of these algorithms were measured and compared with one another. The result showed that the DIC algorithms could achieve the best accuracy performance than SIFT and SURF to process in-plan transformation and deformation scenarios.

In terms of the efficiency, SURF reveals the superiority among the three algorithms. It can be treated to be real time or near real time in processing the experimental image pairs. The DIC algorithm is inefficient compared to SIFT and SURF algorithms.

In general, the DIC algorithm can perform better in measurement accuracy than SIFT and SURF, which reveals that the DIC algorithm is quite potential to be utilized to build the temporary structure monitoring system. However, the operation efficiency of DIC is an obstacle to apply this algorithm into the monitoring system. Therefore, in the future work, this problem will be further investigated by reducing the computing complexity of the algorithm.

This study primarily focused on identifying a superior visual sensing method for monitoring the deflection of temporary structures. Because the failure of a structure can be associated with large deflection or deformation on its load-bearing structural members, promptly obtaining the deflection or deformation will help assess whether the structure is at risk. Such informative observation is particularly important for situations such as cement pouring workers who stand on a temporary structure to conduct their work and other workers working on temporary structures such as scaffolding and formwork support.

Nevertheless, there is still room for improvements of the current study. One is the establishment of the relationship between the structural deformation/deflection and probability of collapse of structure. The change of the magnitude of deformations or deflections over time might be a good indicator for estimation. However, more investigations are needed.

**Acknowledgments** The authors acknowledge the financial support from the West Virginia University Research and Scholarship Grant (Award No.: R-14-037) and National Natural Science Foundation of China (Grant No.: 41302217).

## References

1. Ratay RT (1987) Temporary structures in construction operations. ASCE, Brooklyn
2. Chen RP, Li ZC, Chen YM, Ou CY, Hu Q, Rao M (2014) Failure investigation at a collapsed deep excavation in very sensitive organic soft clay. *J Perform Constr Facil*. doi:[10.1061/\(ASCE\)CF.1943-5509.0000557](https://doi.org/10.1061/(ASCE)CF.1943-5509.0000557)
3. Holt ASJ (2008) Principles of construction safety. Wiley, New York, pp 58–59, 161–163
4. Chintalapudi K, Fu T, Paek J, Kothari N, Rangwala S, Caffrey J, Govindan R, Johnson E, Masri S (2006) Monitoring civil structures with a wireless sensor network. *Internet Comput IEEE* 10(2):26–34
5. Zhu HH, Ho ANL, Yin JH, Sun HW, Pei HF, Hong CY (2012) An optical fibre monitoring system for evaluating the performance of a soil nailed slope. *Smart Struct Syst* 9(5):393–410
6. Ran L, Ye XW, Zhu HH (2011) Long-term monitoring and safety evaluation of a metro station during deep excavation. *Proc Eng* 14:785–792
7. Chen J, Jin G, Meng L (2007) Applications of digital correlation method to structure inspection. *Tsinghua Sci Technol* 12(3):237–243
8. Toole T Michael (2002) Construction site safety roles. *J Constr Eng Manag* 128(3):203–210
9. Paek J, Chintalapudi K, Caffrey J, Govindan R, Masri S (2005) A wireless sensor network for structural health monitoring: performance and experience. In: Proceedings of 2nd IEEE workshop on embedded network sensors, pp 1–10
10. Bennett PJ, Soga K, Wassell IJ, Fidler P, Abe K, Kobayashi Y, Vanicek M (2010) Wireless sensor networks for underground railway applications: case studies in Prague and London. *Smart Struct Syst* 6(5–6):619–639
11. Lynch JP, Loh KJ (2006) A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock Vib Dig* 38(2):91–130
12. Park HS, Lee HM, Adeli H, Lee I (2007) A new approach for health monitoring of structures: terrestrial laser scanning. *Comput Aided Civ Infrastruct Eng* 22(1):19–30
13. Maas H-G, Hampel U (2006) Photogrammetric techniques in civil engineering material testing and structure monitoring. *Photogramm Eng Remote Sens* 72(1):39–45
14. Wahbeh A Mazen, Caffrey JP, Masri SF (2003) A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Mater Struct* 12(5):785
15. Wang Yu, Cuitiño AM (2002) Full-field measurements of heterogeneous deformation patterns on polymeric foams using digital image correlation. *Int J Solids Struct* 39(13):3777–3796
16. Küntz M, Jolin M, Bastien J, Perez F, Hild F (2006) Digital image correlation analysis of crack behavior in a reinforced

- concrete beam during a load test. *Can J Civ Eng* 33(11):1418–1425
17. Govender N (2009) Evaluation of feature detection algorithms for structure from motion. In: *Proceedings of 3rd robotics and mechatronics symposium*, Pretoria
  18. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
  19. Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. *Computer vision—ECCV 2006*. In: *Proceedings of 9th European conference on computer vision*, pp 404–417
  20. Juan L, Gwon O (2009) A comparison of SIFT, PCA-SIFT and SURF. *Int J Image Process* 3(4):143–152
  21. Mikolajczyk K, Schmid C (2005) A performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell* 27(10):1615–1630
  22. Fathi H, Dai F, Lourakis M (2015) Automated as-built 3D reconstruction of civil infrastructure using computer vision: achievements, opportunities, and challenges. *Adv Eng Inf* 29(2):149–161
  23. Segundo MP, Gomes L, Bellon ORP, Silva L (2012) Automating 3D reconstruction pipeline by surf-based alignment. In: *Proceedings of 19th IEEE international conference on image processing*, pp 1761–1764
  24. Cooreman S, Lecompte D, Sol H, Vantomme J, Debruyne D (2008) Identification of mechanical material behavior through inverse modeling and DIC. *Exp Mech* 48(4):421–433
  25. Khan NY, Brendan M, Geoff W (2011) SIFT and SURF performance evaluation against various image deformations on benchmark dataset. In: *Proceedings of 2011 international conference on digital image computing techniques and applications*, Seville
  26. Nghiem AT, Bremond F, Thonnat M, Valentin V (2007) ETI-SEO, performance evaluation for video surveillance systems. In: *IEEE conference on advanced video and signal based surveillance*, London, pp 476–481
  27. Porwik P, Agnieszka L (2004) The Haar-wavelet transform in digital image processing: its status and achievements. *Mach Graph Vis* 13(1/2):79–98
  28. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). *Comput Vis Image Underst* 110(3):346–359
  29. Ta DN, Chen WC, Gelfand N, Pulli K (2009) SURFTRAC: efficient tracking and continuous object recognition using local feature descriptors. In: *IEEE conference on computer vision and pattern recognition*, pp 2937–2944
  30. Bornert M, Brémand F, Doumalin P, Dupré JC, Fazzini M, Grédiac M, Wattrisse B (2009) Assessment of digital image correlation measurement errors: methodology and results. *Exp Mech* 49(3):353–370
  31. Zhao JQ, Zeng P, Lei LP, Ma Y (2012) Initial guess by improved population-based intelligent algorithms for large inter-frame deformation measurement using digital image correlation. *Opt Lasers Eng* 50(3):473–490
  32. Taniguchi A, Furukawa A, Kanasaki S, Tateyama T, Chen YW (2013) Automated assessment of small bowel motility function based on three-dimensional zero-mean normalized cross correlation. In: *Proceedings of 6th international conference on biomedical engineering and informatics*, pp 802–805