Mechanical Systems and Signal Processing 66-67 (2016) 425–436

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Dynamic displacement measurement of large-scale structures based on the Lucas–Kanade template tracking algorithm

Jie Guo*, Chang'an Zhu

Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China (USTC), Hefei, AnHui 230027, PR China

ARTICLE INFO

Article history:
Received 30 July 2014
Received in revised form
1 December 2014
Accepted 11 June 2015
Available online 15 July 2015

Keywords:
Displacement measurement
High-speed camera system
Template tracking
Lucas–Kanade algorithm

ABSTRACT

The development of optics and computer technologies enables the application of the vision-based technique that uses digital cameras to the displacement measurement of large-scale structures. Compared with traditional contact measurements, vision-based technique allows for remote measurement, has a non-intrusive characteristic, and does not necessitate mass introduction. In this study, a high-speed camera system is developed to complete the displacement measurement in real time. The system consists of a high-speed camera and a notebook computer. The high-speed camera can capture images at a speed of hundreds of frames per second. To process the captured images in computer, the Lucas–Kanade template tracking algorithm in the field of computer vision is introduced. Additionally, a modified inverse compositional algorithm is proposed to reduce the computing time of the original algorithm and improve the efficiency further. The modified algorithm can rapidly accomplish one displacement extraction within 1 ms without having to install any pre-designed target panel onto the structures in advance. The accuracy and the efficiency of the system in the remote measurement of dynamic displacement are demonstrated in the experiments on motion platform and sound barrier on suspension viaduct. Experimental results show that the proposed algorithm can extract accurate displacement signal and accomplish the vibration measurement of large-scale structures.

1. Introduction

Large-scale structures, including bridges and buildings, are exposed to various external loads, such as traffic, earthquakes, and gusts during their lifetime. These external loads may induce structural damage that results in costly repair and loss of human lives. Therefore, a system of monitoring the behavior of large-scale structures is necessary to obtain valuable information for structural safety evaluation. Conventional sensors like accelerometers [1,2] are widely used in the industry to study the behavior of large-scale structures. However, conventional sensors can only obtain vibration acceleration information, which does not provide an intuitionistic exhibition of the actual vibration. Several limitations also restrict the application range of this contact-type measurement. For example, the introduction of sensors may change the behavior of structures. The installation of conventional sensors is sometimes difficult, such as installation on a bridge spanning a wide

* Corresponding author. Tel.: +86 15955138765.
E-mail addresses: guojiegj@mail.ustc.edu.cn (J. Guo), changan@ustc.edu.cn (C. Zhu).

http://dx.doi.org/10.1016/j.ymssp.2015.06.004
0888-3270/

river. In many situations, the measurement results may be interfered with by factors such as remote measurement, high temperature, and magnetic fields. For this reason, interest is growing in the development of non-contact displacement measurement techniques, such as speckle photography [3], global positioning systems (GPS) [4,5], and Laser Doppler vibrometer [6]. However, the high costs of these non-contact systems prevent their wide applications.

Optical devices and cameras offer effective alternatives to non-contact measurement with their advantages of remote measurement and no mass introduction. With these devices, what you see is what you get and movement information can be shown visually. Owing to the wide availability of affordable high-quality digital imaging sensors and high-performance computers, cheaper cameras with high resolution have found growing applications in several areas, including engineering monitoring [7–9], human motion [10,11] and underwater measurement [12]. Recently, vision-based techniques were successfully used for the displacement measurement of large-scale structures with good results [13–18]. Wahbeh et al. [13] realized the measurement of displacements and rotations of the Vincent Thomas Bridge in California by using a highly accurate camera in conjunction with a laser tracking reference. Fukuda et al. [16] proposed a camera-based sensor system that uses a robust object search algorithm to measure the dynamic displacements of large-scale structures. Digital image correlation (DIC) [17,18] is also applied in experimental mechanics for non-contact, full-field deformation measurement.

The traditional camera system for displacement measurement is composed of commercial digital cameras and a notebook computer, both of which are now very affordable. However, commercial digital cameras have low frame rates, for example, 30 frames per second (fps), which limit their range of applications to vibration frequencies of less than 15 Hz [14]. Fortunately, numerous high-speed vision systems that can capture images at 1000 fps or even faster have been developed to overcome these restrictions, and several applications using these systems have been reported in [9,19]. In the present paper, a high-speed camera system composed of a notebook computer connected to a camera head with telescopic lenses is developed for displacement measurement. The camera head uses CCD as an image receiver and captures grayscale images at high speed. The pictures are then streamed into the notebook computer through a USB 3.0 interface. On the notebook computer, the captured video can be processed by a software to realize the tracking of an object and extract motion information.

High-speed cameras have high requirements for motion-tracking algorithms. Conventional image-processing techniques are widely used, but they are complex and cumbersome. They require a series of image-processing steps before relevant information can be extracted for practical purposes. These steps include texture recognition, projection of the captured image, and calculation of the actual displacement using target geometry, in which many parameters need to be adjusted simultaneously to accommodate the specific type of image conditioning. As a result, the computational efficiency and the robustness of the computer algorithms are negatively affected, and expertise is required to adjust the processing parameters for optimal performance. In addition, conventional processing systems previously developed require the installation of a target panel on the measurement point of the structures in advance [13–15]. The target panel with a pre-designed high-contrast black and white pattern contributes to the high-precision tracking of the target movement. However, the installation of the target panel is sometimes rather difficult, such as a bridge spanning a wide river. Although several applications without the targets were developed recently [16,20–22], their measurable locations on the structures are limited by specific characteristics such as cables and gusset plates. In this study, the Lucas–Kanade template tracking algorithm in the field of computer vision is introduced to achieve displacement measurement by using existing features such as edges and texture on the structures instead of an extra target panel. Compared with the paper [20] in which Lucas–Kanade optical flow method was used to obtain the direction of vibration and the canny edge detection algorithm was applied to extract 1D motion displacement, the proposed method got rid of additional image-processing techniques with less adjustable parameters.

Template tracking is a widely studied algorithm in computer vision. It belongs to the direct type of region-based tracking method [23,24] and can date back to Lucas and Kanade [25]. A unifying framework for the Lucas–Kanade algorithm has been reported in [26], and its goal is to minimize the sum of squared error between the template and the image warped back onto the coordinate frame of the template. In particular, the inverse compositional (IC) algorithm [27] is a more efficient version, where the roles of the template and the image are switched, and the Hessian need not be updated at each iteration. Through the direct use of region content of the image, the task of object tracking is accomplished by extracting the template region in the first frame and finding the most matching region in the following frames. In this way, the moving object can be tracked and the motion information can be extracted. Template-tracking algorithms carry both spatial and appearance information, and thus perform well with high levels of robustness and accuracy. They have been widely applied in various applications, including object tracking [28,29], parametric and motion estimation [30,31], and medical image registration [32,33].

Given its excellent characteristics, the Lucas–Kanade IC algorithm is introduced in this study on displacement measurement to solve the problems of conventional image-processing techniques. A rapid and real-time modification for translation transformation based on the original IC algorithm is also proposed. After several optimizations, the computation time of one extraction in the modified algorithm is reduced to less than 1 ms, thereby providing the potential of real-time extraction to the high-speed captured video. Finally, the proposed algorithm is utilized in the high-speed camera system to realize the displacement measurement of actual large-scale structures. And two experiments are carried out under laboratory and realistic conditions for performance verification. The positive results demonstrated the accuracy and efficiency of the camera system in the remote measurement of dynamic displacement.

This paper is organized as follows: Section 2 introduces the components and capability parameters of the high-speed camera system. Section 3 presents the theory of Lucas–Kanade template tracking algorithm. Section 4 describes the

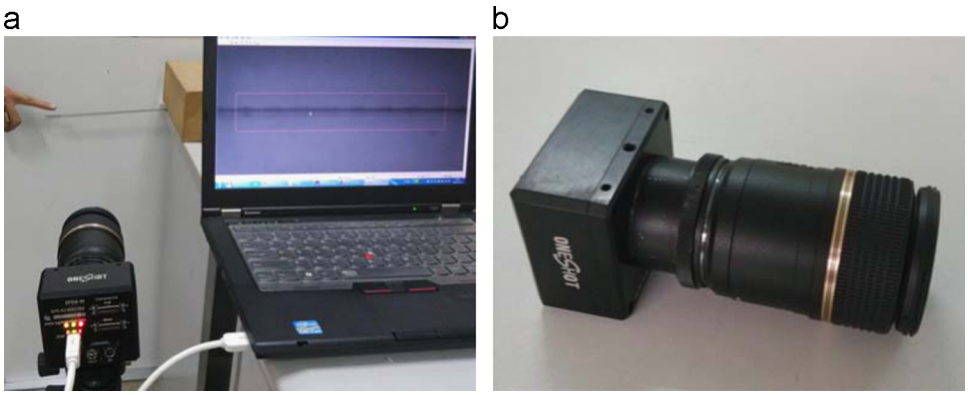

Fig. 1. High-speed camera system: (a) the experimental setup and (b) the high-speed camera head.

modified IC algorithm for translation transformation and provides a simulation example. Section 5 presents two experiments for performance verification. Section 6 concludes.

2. High-speed camera system

The high-speed camera system developed in this study, as shown in Fig. 1(a), consists of a notebook computer (Intel Core processor 2.9 G Hz; 2.75 GB RAM) connected to a camera head with telescopic lenses. The telescopic lenses with large optical zooming capability in Fig. 1(b) can be adjusted appropriately to perceive structural motion at different distances. The camera head uses a CCD sensor as image receiver and can capture 8-bit grayscale images at high speed, which are streamed into the notebook computer through USB 3.0 interface. The image-processing software on the notebook computer can use the template-tracking algorithm to track a particular fast-moving object and extract motion information. The camera frame rate can reach 1000 fps when the image resolution is set as 300 pixels × 300 pixels.

To measure the displacement of remote structures, a target panel may be installed in advance or specific features are selected on the measuring location of the structure. Then the camera system is ready to capture images from a remote location, and the displacement time history of the structure can be obtained by applying the template-tracking algorithm to the digital video images.

3. Theory of the Lucas–Kanade algorithm

In the rest of this paper, we follow the notations of Baker and Matthews [26] as closely as possible. Let $I_i(\mathbf{x})$ stand for the nth image in a given video sequence, where $\mathbf{x}=(x,y)^T$ is a column vector containing the pixel coordinates and $i=0,1,2,\ldots$ is the frame number. The template $T(\mathbf{x})$ is an extracted subregion of the initial frame I_0, which contains existing features to be tracked in the image sequence. Thus, the goal of the template tracking algorithm is to align the template image $T(\mathbf{x})$ to the subsequent frames $I_i(\mathbf{x})$ to track the object motion.

The alignment of images in a sequence can be parameterized as a warp function $\mathbf{W}(\mathbf{x};\mathbf{p})$. This function transforms the pixel coordinate $\mathbf{x}$ in the coordinate frame of the template T to the sub-pixel coordinate $\mathbf{W}(\mathbf{x};\mathbf{p})$ in the coordinate frame of the subsequent image I_i. $\mathbf{p}=(p_1,\ldots,p_n)^T$ denotes a vector of transformation parameters of the warp function. The objective of template-based tracking is to search for the optimized parameters $\mathbf{p}$ so that the warped image $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$ matches with the template $T(\mathbf{x})$. In the original Lucas–Kanade algorithm, the objective can be solved by minimizing the sum of square errors between $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$ and $T(\mathbf{x})$ as

$$\sum_{\mathbf{x}\in T}[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))-T(\mathbf{x})]^2 \tag{1}$$

where the summation is over all pixels of the template $T(\mathbf{x})$ and the minimization is performed with respect to $\mathbf{p}$. Note that the computation of $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$ requires interpolating the image I_i at the sub-pixel locations $\mathbf{W}(\mathbf{x};\mathbf{p})$.

Owing to its non-linear nature, the optimization problem of Eq. (1) is done iteratively solving for increments $\Delta\mathbf{p}$ to update the current parameters $\mathbf{p}$. Therefore, the error function (1) can be rewritten as

$$\sum_{\mathbf{x}\in T}[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}+\Delta\mathbf{p}))-T(\mathbf{x})]^2 \tag{2}$$

with respect to $\Delta\mathbf{p}$, and the parameters are updated as follows:

$$\mathbf{p}\leftarrow\mathbf{p}+\Delta\mathbf{p} \tag{3}$$

These two steps are iterated until the estimates of $\mathbf{p}$ converge. Typically, the test for convergence is whether the norm of the vector $\Delta\mathbf{p}$ is below a threshold ϵ (i.e., $\|\Delta\mathbf{p}\|\le\epsilon$).

The optimization problem of Eq. (2) is a composite function. To obtain a closed-form solution, the first-order Taylor expansion is used to approximate the composite function $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}+\Delta\mathbf{p})$ as

$$I_i\left(\mathbf{W}\left(\mathbf{x};\mathbf{p}+\Delta\mathbf{p}\right)\right)=I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))+\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\Delta\mathbf{p} \tag{4}$$

where $\nabla I_i=\left(\frac{\partial I_i}{\partial x},\frac{\partial I_i}{\partial y}\right)$ is the gradient of image I_i evaluated when the current warp function $\mathbf{W}(\mathbf{x};\mathbf{p})$ is applied, and the term $\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$ represents the Jacobian of the warp.

Therefore, the error function (2) can be written as

$$\sum_{\mathbf{x}\in T}\left[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))+\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\Delta\mathbf{p}-T(\mathbf{x})\right]^2 \tag{5}$$

Minimizing the error function of Eq. (5) is a least-squares problem. The closed-form solution can be obtained by taking the partial derivative first and then setting it to zero. Then the solution can be solved as [26]

$$\Delta\mathbf{p}=H^{-1}\sum_{\mathbf{x}\in T}\left[\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T[T(\mathbf{x})-I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))] \tag{6}$$

where H is the Gauss–Newton approximation to the $n\times n$ Hessian matrix:

$$H=\sum_{\mathbf{x}\in T}\left[\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T\left[\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right] \tag{7}$$

The Lucas–Kanade algorithm then switches into iteratively applying Eqs. (6) and (3). The entire procedure is summarized in Table 1.

3.1. Inverse compositional algorithm

As several studies have pointed out [26,27], the original Lucas–Kanade algorithm suffers from huge computational cost in re-evaluating the Hessian in every iteration. Fortunately, an efficient version, the IC algorithm, promises to reduce the computation without any significant loss of efficiency, in which the roles of the template and the image are switched. In this case, the following expression is to be iteratively minimized:

$$\sum_{\mathbf{x}\in T}[T(\mathbf{W}(\mathbf{x};\Delta\mathbf{p}))-I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))]^2 \tag{8}$$

with respect to $\Delta\mathbf{p}$ and then updates the warp

$$\mathbf{W}(\mathbf{x};\mathbf{p})\leftarrow\mathbf{W}(\mathbf{x};\mathbf{p})\bigcirc\mathbf{W}(\mathbf{x};\Delta\mathbf{p})^{-1} \tag{9}$$

Table 1
The Lucas–Kanade template tracking algorithms.

The Lucas–Kanade algorithm	The inverse compositional algorithm	The modified IC algorithm for translation
Iterate:	**Pre-compute**:	**Pre-compute**:
(1) Warp I_i with $\mathbf{W}(\mathbf{x};\mathbf{p})$ to compute $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$	(3) Evaluate the gradient ∇T of the template $T(\mathbf{x})$	(3) Evaluate the gradient ∇T of the template $T(\mathbf{x})$
(2) Compute the error image $T(\mathbf{x})-I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$	(4) Evaluate the Jacobian $\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$ at $(\mathbf{x};\mathbf{0})$	(6) Compute the Hessian matrix using Eq. (16)
(3) Warp the gradient ∇I_i with $\mathbf{W}(\mathbf{x};\mathbf{p})$	(5) Compute $\nabla T\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$	
(4) Evaluate the Jacobian $\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$ at $(\mathbf{x};\mathbf{p})$	(6) Compute the Hessian matrix using Eq. (12)	**Iterate:**
(5) Compute $\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$		(0) Round off the parameter: $\mathbf{p}_r\leftarrow round(\mathbf{p})$
(6) Compute the Hessian matrix using Eq. (7)	**Iterate:**	(1) Cut $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}_r))$ from I_i
(7) Compute $\sum_{\mathbf{x}\in T}\left[\nabla I_i\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T[T(\mathbf{x})-I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))]$	(1) Warp I_i with $\mathbf{W}(\mathbf{x};\mathbf{p})$ to compute $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$	(2) Compute the error image $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))-T(\mathbf{x})$
(8) Compute $\Delta\mathbf{p}$ using Eq. (6)	(2) Compute the error image $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}_r))-T(\mathbf{x})$	(7) Compute $\sum_{\mathbf{x}\in T}(\nabla T)^T[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}_r))-T(\mathbf{x})]$
(9) Update the parameter $\mathbf{p}\leftarrow\mathbf{p}+\Delta\mathbf{p}$	(7) Compute $\sum_{\mathbf{x}\in T}\left[\nabla T\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))-T(\mathbf{x})]$	(8) Compute $\Delta\mathbf{p}$ using Eq. (15)
until $\|\Delta\mathbf{p}\|\le\epsilon$	(8) Compute $\Delta\mathbf{p}$ using Eq. (11)	(9) Update the parameter $\mathbf{p}\leftarrow\mathbf{p}-\Delta\mathbf{p}$
	(9) Update the parameter $\mathbf{W}(\mathbf{x};\mathbf{p})\leftarrow\mathbf{W}(\mathbf{x};\mathbf{p})\bigcirc\mathbf{W}(\mathbf{x};\Delta\mathbf{p})^{-1}$	until all $\|\Delta\mathbf{p}\|\le 0.5$.
	until $\|\Delta\mathbf{p}\|\le\epsilon$.	

Performing a first-order Taylor expansion on Eq. (8) gives

$$\sum_{\mathbf{x}\in T}\left[T(\mathbf{W}(\mathbf{x};\mathbf{0}))+\nabla\mathbf{T}\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\Delta\mathbf{p}-\mathbf{I}_i(\mathbf{W}(\mathbf{x};\mathbf{p}))\right]^2 \tag{10}$$

Without loss of generality, assuming that $\mathbf{W}(\mathbf{x};\mathbf{0})$ is the identity warp, the solution to this least-squares problem is

$$\Delta\mathbf{p}=H^{-1}\sum_{\mathbf{x}\in T}\left[\nabla T\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))-T(\mathbf{x})] \tag{11}$$

where H is the Hessian matrix

$$H=\sum_{\mathbf{x}\in T}\left[\nabla T\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right]^T\left[\nabla T\frac{\partial\mathbf{W}}{\partial\mathbf{p}}\right] \tag{12}$$

where the Jacobian $\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$ is evaluated at $(\mathbf{x};\mathbf{0})$; the gradient of T, rather than the gradient of I_i, is constant. Thus, H does not depend on the warping parameters and can be pre-computed, which greatly improves the efficiency of the algorithm. The IC algorithm then switches into iteratively applying Eqs. (11) and (9). The entire procedure is summarized in Table 1.

As [26] pointed out, two algorithms empirically perform equivalently in convergence properties, but the IC algorithm is far more computationally efficient than the original Lucas–Kanade algorithms. Because the most time consuming in Step 6, the computation of the Hessian, is performed once as a pre-computation, the computation cost of each iteration reduces from $O(n^2N+n^3)$ to $O(nN+n^2)$, where n is the number of warp parameters and N is the number of pixels in the template T. Two algorithms show different responses to noise. If the template has more noise, the original Lucas–Kanade algorithm should be used, whereas the IC algorithm is more suitable if the input image has more noise. Efficiency is the foremost concern in this paper, and thus the IC algorithm is used.

4. Real-time inverse compositional algorithm for translation

In this section we introduce the IC algorithm to the field of displacement measurement and propose a simple, efficient modification to meet the requirements of real-time measurement in the high-speed camera system.

4.1. Inverse compositional algorithm for translation

In the original algorithm, the warp function plays an important role in tracking different types of motion. For example, 2D affine transformation is widely used in common tracking. It consists of six independent parameters and takes into account many linear transformations, including translation, rotation, shear mapping, and scaling. To put it simply, the displacement measurement presupposes that the movements of structures are mostly in-plane motions with a limited range, and deformation and rotation can be neglected in a short capturing time. Therefore, translation transformation is utilized instead as the warp function $\mathbf{W}(\mathbf{x};\mathbf{p})$, which consists of two independent parameters $\mathbf{p}=(p_x,p_y)^T$

$$\mathbf{W}(\mathbf{x};\mathbf{p})=\begin{pmatrix}x+p_x\\ y+p_y\end{pmatrix} \tag{13}$$

where p_x and p_y represent the displacement in the horizontal and vertical directions respectively.

Based on the concise form of translation transformation, the procedure of IC algorithm can be simplified significantly. First, the Jacobian of the warp, $\frac{\partial\mathbf{W}}{\partial\mathbf{p}}$, degenerates into an identity matrix

$$\frac{\partial\mathbf{W}}{\partial\mathbf{p}}=\begin{pmatrix}\frac{\partial W_x}{\partial p_x} & \frac{\partial W_x}{\partial p_y}\\ \frac{\partial W_y}{\partial p_x} & \frac{\partial W_y}{\partial p_y}\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{14}$$

which makes the update of $\Delta\mathbf{p}$ in Eq. (11) more concise as

$$\Delta\mathbf{p}=H^{-1}\sum_{\mathbf{x}\in T}(\nabla T)^T[I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))-T(\mathbf{x})] \tag{15}$$

and the Hessian matrix H takes the form

$$H=\sum_{\mathbf{x}\in T}(\nabla T)^T\nabla T \tag{16}$$

The Hessian matrix with its inverse matrix can be calculated easily because the warp function only has two parameters and the size of the Hessian matrix is 2×2.

Moreover, the warp update eliminates the warp inversion in Eq. (9) and is simplified as

$$\mathbf{p}\leftarrow\mathbf{p}-\Delta\mathbf{p} \tag{17}$$

4.2. Modified inverse compositional algorithm

The above simplifications make the expression of the IC algorithm concise and easy to understand. Other modifications are further proposed to improve the algorithm efficiency. In the iteration steps of IC algorithm, two steps need to be emphasized. First, in Step 1, the computation of $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}))$ requires interpolating the image I_i at the sub-pixel positions $\mathbf{W}(\mathbf{x};\mathbf{p})$, which sweeps on the whole image and calculates the estimation value of each pixel according to the surrounding pixels. Interpolation is usually very time consuming, especially if a complex interpolation type is selected. Second, the termination of the iteration process depends on the convergence threshold ϵ, i.e., $\|\Delta\mathbf{p}\| \le \epsilon$, which is the only parameter specified by humans in advance. A small threshold may improve the accuracy, but additional iterations and low efficiency must be introduced. Thus, the selection of the threshold requires a trade-off between accuracy and efficiency.

To solve these problems, we introduce one additional operation before Step 1, that is, rounding $\mathbf{p}$ off to the nearest integer denoted as $\mathbf{p}_r = (p_{xr}, p_{yr})^T$. This operation can avoid image interpolation and obtain $I_i(\mathbf{W}(\mathbf{x};\mathbf{p}_r))$ by a simple cutting operation from I_i, as shown below

$$I_i(\mathbf{W}(\mathbf{x};\mathbf{p}_r)) = I_i(x+p_{xr}, y+p_{yr}) \tag{18}$$

where x and y are integers in the range $[1,w]$ and $[1,h]$; and w and h are the width and height of the template $T(\mathbf{x})$. This modification is so simple that hardly any computation is introduced. In this case, the termination condition can be simplified as $|\Delta\mathbf{p}| \le 0.5$, where $|\Delta p_x| \le 0.5$ and $|\Delta p_y| \le 0.5$ because the parameter update greater than 0.5 will be omitted due to the rounding-off operation. This improvement automates the algorithm further, and no additional parameters need to be specified. However, it may reduce calculation accuracy. In the original IC algorithm, the sub-pixel iteration, which is when $\Delta\mathbf{p} < 1$, can be performed many times until $\Delta\mathbf{p}$ is less than a small threshold ϵ, but the rounding-off operation in the modified algorithm ends the sub-pixel iteration at $\Delta\mathbf{p} < 0.5$. Decreasing the time consumed for interpolation operation sacrifices the accuracy of the modified algorithm to a certain extent, demonstrating that efficiency improves at the cost of accuracy. Fortunately, the modified algorithm can still obtain sub-pixel results.

Based on the analysis above, the modified IC tracking algorithm is simple and efficient but still provides sub-pixel accuracy. Its procedure is summarized in Table 1. As shown in the table, each algorithm iteration involves a numerical rounding-off (Step 0), an image cutting (Step 1), an image difference (Step 2), a collection of image "dot-products" (Step 7), a multiplication with the inversion of Hessian (Step 8), and a parameter update (Step 9). All the operations are easy to perform except for Step 7. In spite of being the most time-consuming step, the collection of image "dot-products" is just a multiplication of two matrices. For typical problems with a template size of 100×100, Step 7 just contains $20,000 (= n \times N = 2 \times 10,000)$ floating-point multiplication. In a typical computer, such as that specified in Section 2, this operation can be performed within 0.1 ms and the computation time of each extraction is within 1 ms (1000 Hz). Therefore, the modified IC algorithm can be used in the high-speed camera system to measure displacement in real time.

4.3. Simulation test

We now present a simple example to show the improving performance of the modified IC algorithm. Consider an image containing a screw, as shown in Fig. 2(a). Screw is an important part which is widely used in engineering and large-scale structures. As reported in [16], when the target panel is difficult to install, specific features or components, such as screws and gusset plates, can be regarded as the tracking object. Here, we take the screw as the template for tracking. The template is marked by a green box (100×100) in Fig. 2(a). The red box indicates the initial warp that shifted 10 pixels upward and 5 pixels toward the right. Based on this initial condition, the template tracking algorithm is implemented to find the exact position.

The entire tracking process of the modified IC algorithm is shown in Fig. 2(b)–(f). The dotted red box represents the new updating position after each iteration, and the current error image between the template and the warped image is attached at the bottom right side. The tracking trajectory is plotted by the red curve.

After 11 iterations, the modified IC algorithm successfully finds the exact position of the screw. As the error images visually illustrate the approaching process to the position, their RMS values against the algorithm iteration number are displayed in Fig. 3(a). These values monotonically decay to zero. The moving coordinate values of the iterations are plotted in Fig. 3(b). For comparison, the original IC algorithm is also performed; its results are displayed as a green curve in Fig. 3. The results reveal that two algorithms converge at nearly similar iteration speed. The modified one runs without any loss in both accuracy and convergence in spite of the simplifications made. In terms of time consumption, each iteration of the modified algorithm takes about 0.1 ms while each iteration of the original algorithm takes about 6 ms even if the simplest linear interpolation is used. Therefore, the modified algorithm is much more computationally effective than the original one because of the significant improvement in time efficiency.

In addition, the shift in each iteration is only about 1 pixel to the right direction. To improve the convergence speed, we can enlarge the step size of each iteration by multiplying it with a factor, which changes along with the iteration to ensure accuracy

$$\Delta\mathbf{p} \leftarrow \Delta\mathbf{p} \times \lambda_i \tag{19}$$

where i is the index of iteration. At the time of the initial iteration, λ_i is large in order to rapidly approach the exact position, but after several iterations, λ_i degenerates to 1 for fine adjustment near the exact position. The red curve in Fig. 3 displays the improving performance if $\lambda_i = 1 + 2 \times 0.8^x$. Nevertheless, this improvement is not a major issue because displacement between

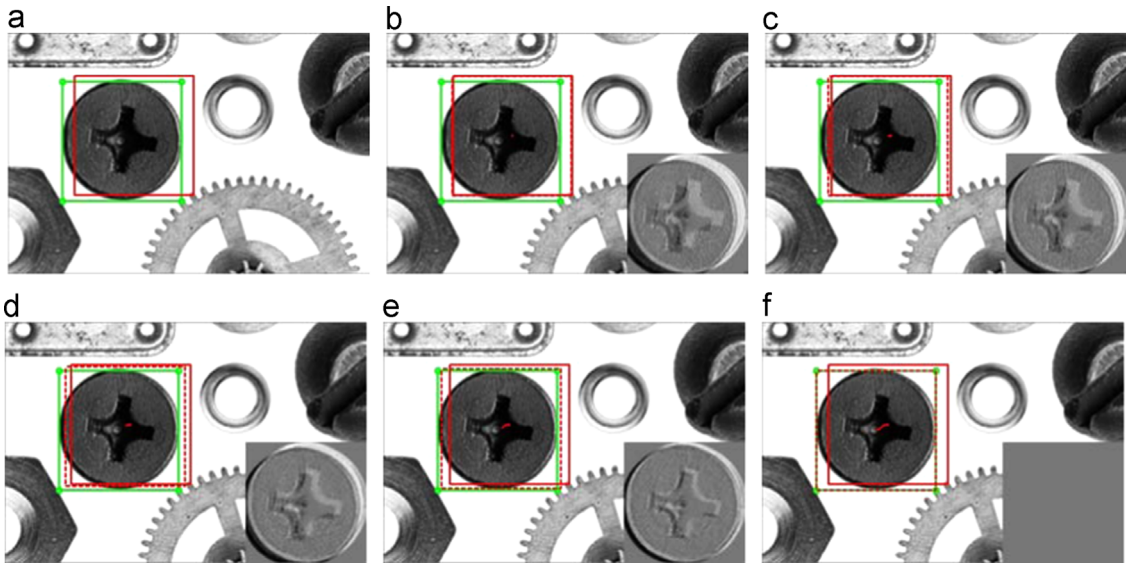

Fig. 2. Example of tracking a screw template: (a) initial image containing a screw, (b)–(f) the iterative tracking process of the modified IC algorithm. The green box is the actual position of the screw, the red box is the initial shifted position where tracking start, and the dotted red box represents the updating position of the iterative process. The small sub-images at the bottom right sides of (b)–(f) are the current error images between the template and the warped image. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

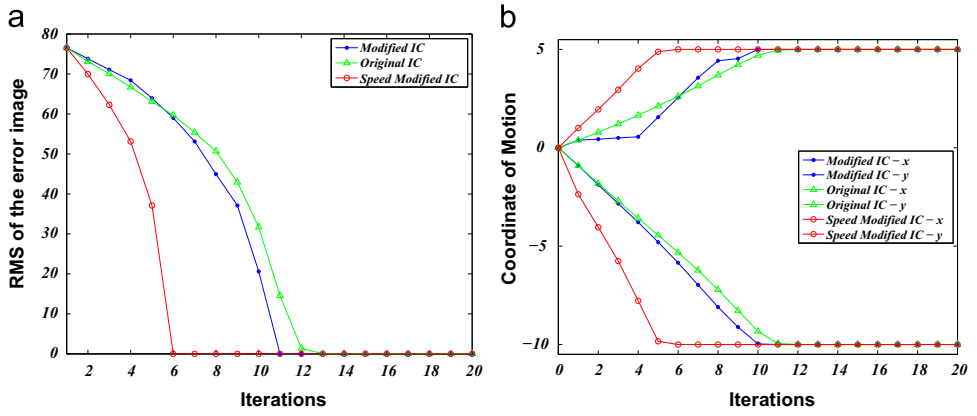

Fig. 3. Comparisons of the RMS error and tracking process among the original IC algorithm, our modified algorithm, and the variable-step improvement. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

two frames in the high-speed camera is very little, for example, less than 2 pixels. Therefore, the tracking algorithm can converge generally within three iterations as verified by the experiments in Section 5.

With its improvement in time efficiency and thus rapid realization of the motion displacement extraction, the modified IC algorithm has been integrated into a real-time measurement software. This software contains several modules. The high-speed camera module can adjust and control the parameters of the digital camera, such as contrast, brightness, and exposure time. The calibration module can compute the actual displacement of 1 pixel based on a target with known size. The image-capturing module can collect streaming image data from the camera for real-time processing by the template-tracking module, in which the modified IC algorithm is performed to extract the displacement of the template. The entire software is implemented based on the Qt and OpenCV libraries so that it can control the camera and handle images effectively. Together with the high-speed camera, the real-time measurement system described in Section 2 is created and can realize the displacement measurement of actual large-scale structures.

5. Experimental verification

5.1. Motion platform experiment

To evaluate the tracking accuracy of the modified IC algorithm, an experimental verification was carried out on a laboratory platform installed with a conventional grating ruler as shown in Fig. 4(a). Grating ruler, also called incremental

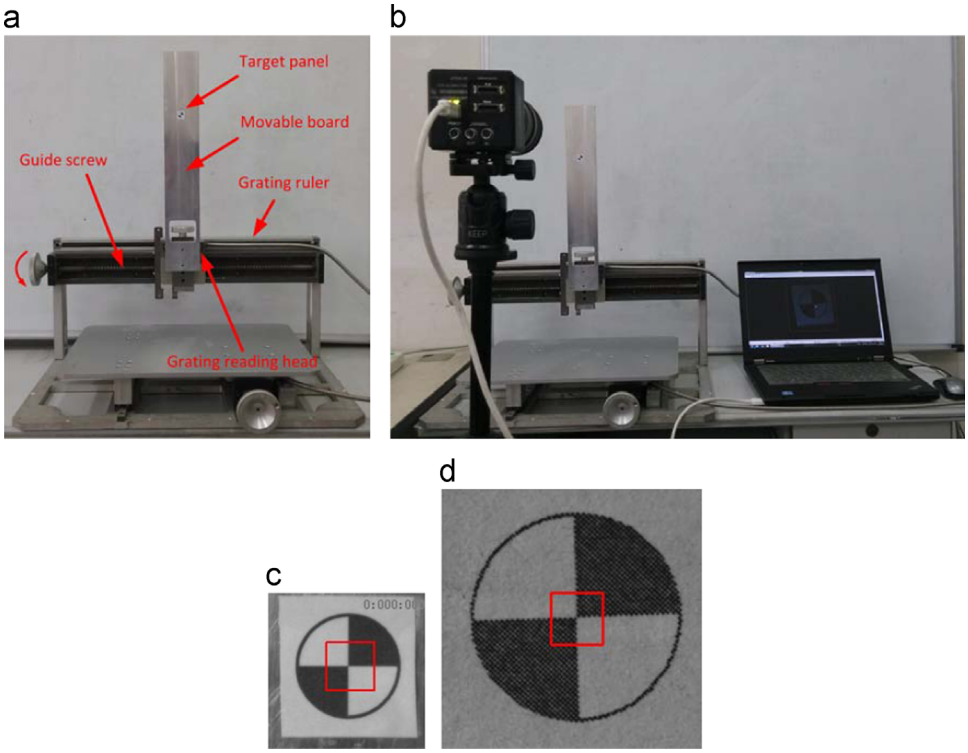

Fig. 4. Motion platform experiment for evaluating the tracking accuracy of the modified IC translation algorithm: (a) the grating ruler and guide screw installed on the motion platform, (b) the experimental setup to accomplish the motion platform experiment, (c) a captured image of the video camera 3 m away from the platform, and (d) another captured image of the video camera 1 m away from the platform. The red box in the captured image represents the position of the template. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

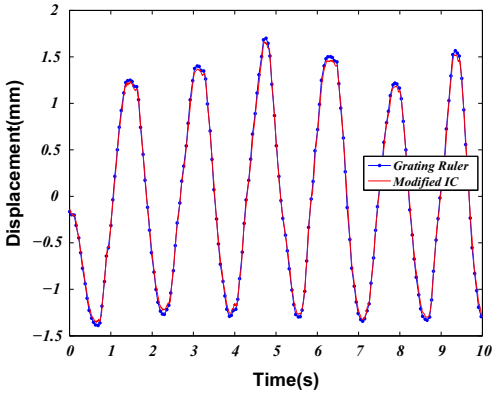

Fig. 5. Comparison of results from the video camera 3 m away from the platform.

grating displacement sensor, is a high accuracy displacement measurement tool that uses the moire fringe technology of grating and photoelectric conversion. It has numerous advantages, such as stability, reliability, and high accuracy. In this experiment, the vision-based high-speed camera system was experimentally evaluated against the grating ruler. As shown in Fig. 4(a), a grating ruler and a guide screw were fixed on the platform. The guide screw can drive the grating reading head together with a movable board to produce horizontal displacement. The displacement can be recorded simultaneously by the grating ruler and the high-speed camera system for comparison. The sampling frequency of the grating ruler is 20 Hz and the grating space is 0.02 mm.

The movable board does not have obvious features, so a pre-designed black and white circular target panel with a 20 mm diameter was fixed on the board. The video camera was placed at a stationary position 3 m away from the platform. Fig. 4(b) shows the experimental setup. The camera system captured videos at 200 fps and extracted the displacement of the target panel by performing the modified IC algorithm on the connected computer. The size of the captured image was 160×160. An image captured by the camera is displayed in Fig. 4(c), in which the red box with the size of 50×50 was selected as the

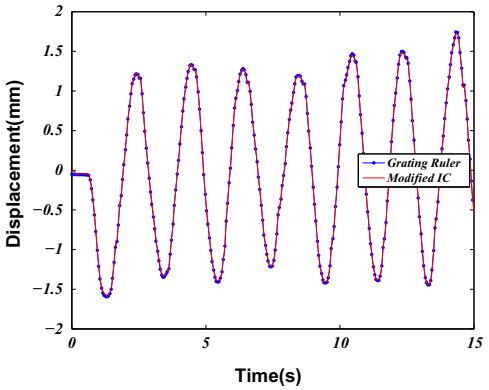

Fig. 6. Comparison of results from the video camera 1 m away from the platform.

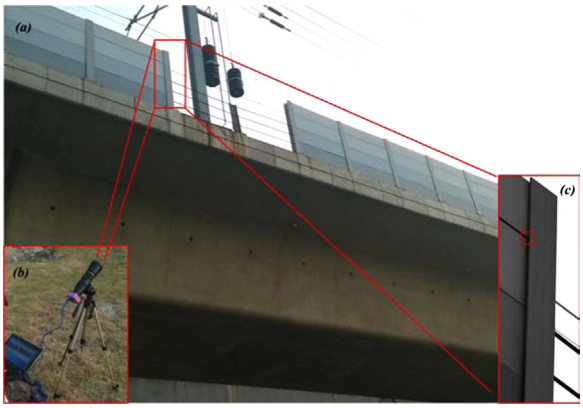

Fig. 7. (a) Viaduct installed with sound barriers, (b) the experimental setup, and (c) an image of the shot video.

template. To measure displacement in real-life dimensions, the actual size of the pre-designed target panel was calculated. Calculations show that 20 mm corresponds to 108 pixels in the captured image; therefore, 1 pixel corresponds to about 0.185 mm.

The guide screw was driven by a manual arbitrary input, and measurement was taken for 10 s. The horizontal displacement time history measured by the vision-based system was compared with that measured by the grating ruler as shown in Fig. 5. The results agree well, with an average tracking error of 0.035 mm, which is relatively small considering the amplitude of about 1.5 mm. Therefore, the modified IC algorithm could accurately track the template's movement, having demonstrated its good trackability to dynamic response. Moreover, if the modified IC algorithm was performed on the computer of Section 2, all the frames converged within three iterations, and more than 93% were performed within two iterations. All the elapsed time were less than 0.3 ms, and more than 90% were completed within 0.1 ms. The excellent result demonstrates the rapidity of the proposed algorithm.

The accuracy of the vision-based tracking system depends on the type of lens and the distance between the camera and the target. In another experiment, the video camera was placed 1 m away from the platform. In this case, the size of the captured image was 500×500 and the pixel size was improved to 0.025 mm/pixel. An image captured by the camera is displayed in Fig. 4(d), in which the red box with the size of 100×100 was selected as the template. Again, the guide screw was driven, and the horizontal displacement time histories measured by the vision-based system and the grating ruler are shown in Fig. 6. Owing to the increase in pixel resolution, the results of the vision-based system and the grating ruler agree better and the average tracking error decreased to 0.016 mm. Therefore, decreasing the distance between camera and target is a simple and effective way to improve the measurement accuracy. Such a technique is, however, not always appropriate, especially for remote measurement. A better lens, like telephoto lens, can be used instead to capture images of the target in a remote distance and obtain a higher resolution. This option was employed in the succeeding experiment on a sound barrier installed on the suspension viaduct.

This platform experiment demonstrates that the vision-based system can measure dynamic displacement with reasonable accuracy, and measurement accuracy and spatial resolution can be improved further by decreasing the measuring distance. Nevertheless, because the modified IC algorithm is a sub-pixel tracking method, its accuracy is related to the pixel resolution but not limited by it.

5.2. Sound barrier experiment

To validate the effectiveness of the proposed algorithms in extracting the motion displacement of actual large-scale structures, an experiment on sound barrier was performed as shown in Fig. 7(a). Sound barrier, also called sound wall or noise barrier, is a platy structure designed to protect inhabitants on both sides of a railway from noise pollution. It is the most effective method of mitigating roadway, railway, and industrial noises. However, with the increase in train speed, strong suction and impact are introduced when high-speed trains pass by the section installed with sound barriers, and the aroused vibration may cause fatigue of materials and loosen assembly. If the sound barriers are damaged and fall on the railway track, disastrous consequences may result. To improve the design of a sound barrier and increase its service life, the vibration status should be measured first at the passing of a high-speed train.

However, installing the traditional measurement devices or attaching a pre-designed target panel to the hard-to-access barrier on the suspension viaduct is difficult. Vision-based measurement using the modified IC algorithm is a suitable alternative to track the existing feature on the barrier and accomplish the non-contact measurement.

The viaduct installed with sound barriers is located in KunShan City, as shown in Fig. 7(a). The high-speed camera system with a telephoto lens was placed on the ground below the viaduct as shown in Fig. 7(b). The distance between the camera and the barrier is about 30 m. When a high-speed train passes by, the camera can record video containing the vibration of the sound barrier at 232 fps. An actual image captured by the camera is displayed in Fig. 7(c), from which some existing seams and edges can be selected as a template for motion tracking. The position of the template, whose size is 50×50, is displayed as the red box in Fig. 7(c). The actual size of 1 pixel is 0.78 mm/pixel, which was calculated by using the known physical size of the barrier. This high image resolution in the case of remote measurement is attributed mainly to the use of telephoto lens. Note that the deformation of the barrier is very small compared to the vibration; thus, the effect on the extracted result may be negligible. The barrier is shot at a certain elevation because of the limited space, so the vibration in the video should be considered as the projection of actual movements on the imaging plane of the camera, which has no effect on the extraction of vibration properties.

The vertical displacement with its corresponding Fourier spectrum after the modified IC extraction algorithm was applied to the barrier video is shown in Fig. 8. The displacement time history clearly displayed the wave shape when the train passed by. The moment of train arrival is marked red in 8(a). In the Fourier spectrum shown in 8(b), three obvious spectral peaks can be observed at 10.42, 21.07, and 45.77 Hz, which can be considered as the characteristic frequencies of the sound barrier. Based on these results, we can grasp the dynamic characteristic of the sound barrier which was difficult to detect before.

Another extraction algorithm called the normalized cross correlation (NCC) was also applied to the measured video images as a post-processing algorithm. The NCC [34], as a template matching technique, is generally used for the DIC and is widely applied to the measurement of image deformation and displacement. In this experiment, it was used to compare with the modified IC algorithm to determine the effectiveness of the latter. The extraction result of NCC, if it used the same template image as did the modified IC algorithm, is shown in Fig. 8. The two methods clearly obtained similar results in terms of the time-domain wave shape and frequency-domain spectrum.

Both algorithms can successfully accomplish the displacement extraction. However, the NCC is accomplished by scanning a template on the whole image to find the best match position, which results in considerable calculation and extensive time spent for one extraction between two frames (exceeds tens of milliseconds). Consequently, the NCC can only be applied to extract displacement as a post-processing algorithm. On the contrary, the modified IC algorithm runs very fast. Same as the experiment done on the grating ruler, every extraction converges within three iterations and more than 98% extractions are performed within two iterations. Moreover, the elapsed time of each extraction is less than 0.6 ms and more than 60% are completed within 0.1 ms. In essence, this rapidity is attributed to the fact that the computation of modified IC algorithm is

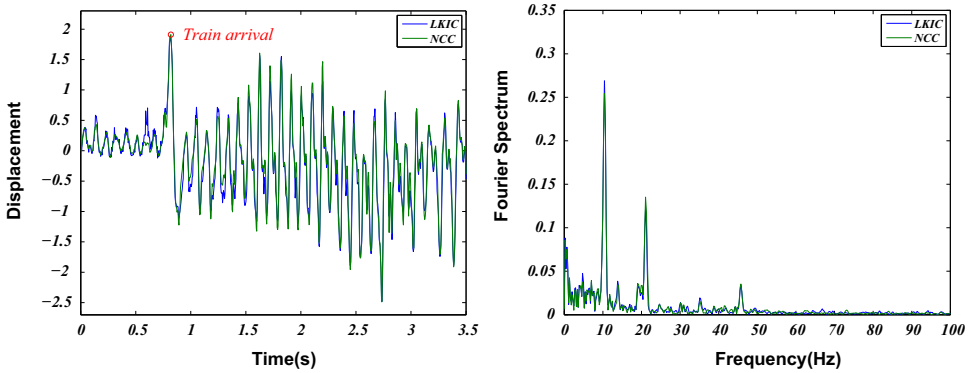

Fig. 8. Extraction results from the application of the modified IC algorithm and the NCC method to the vibration video of the sound barrier: (a) the extracting displacement time history and (b) the Fourier spectrum of the displacement. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

mainly composed of a simple multiplication of two matrices, whereas the NCC requires cumbersome template matching. The NCC also necessitates an extra estimating step to obtain sub-pixel accuracy, whereas the modified IC algorithm is inherently a sub-pixel algorithm.

The sound barrier experiment demonstrates that the modified IC algorithm is an effective vision-based method because of its high-efficiency and high-accuracy advantages. The present high-speed displacement measurement system has the ability to monitor the dynamic response of actual large-scale structures in real time, even in the case of vibration frequency greater than 15 Hz which is impossible for ordinary camera with low frame rate(30 fps).

6. Conclusions

In this study, optical cameras were used for remote measurement of dynamic displacement of large-scale structures and a high-speed camera system was developed for online measurement. To meet the real-time requirements, the Lucas–Kanade template tracking algorithm in the field of computer vision was introduced and a modified IC algorithm for translation transformation was proposed. Through the optimization of the original IC algorithm, the interpolation step was replaced by a rounding-off operation and an image cutting. The computation of one iteration mainly comprised a collection of image "dot-products", i.e., a multiplication of two matrices, that could be performed rapidly within 0.1 ms even in a common computer without GPU acceleration. The great reduction in the processing time makes the proposed algorithm be smoothly executed on the high-speed camera system in real-time. Besides, by implementing the modified IC algorithm, the motion displacement of hard-to-access structures could also be measured through tracking existing features on the structures without requiring the installation of an extra target panel. This may be considered a very meaningful contribution to the non-contact and remote measurement of hard-to-access structures.

Two experimental studies validated the performance of the proposed algorithm. The motion platform experiment demonstrated the accuracy of the dynamic displacement measurement by comparing the results of the high-speed camera system with those of a conventional sensor of a grating ruler. An experiment on a sound barrier on suspension viaduct further displayed the reliability to accurately extract vibrating signals of real-life structures. The results using the proposed algorithm achieved excellent agreement with those of another post-processing algorithm, namely NCC.

The changes in shading, lighting, and background conditions in the field made it difficult to extract accurate displacement from video images using the Lucas–Kanade algorithm, further work will be carried out to improve the robustness of the modified IC algorithm to brightness variations.

Acknowledgments

This work was supported by the National Key Basic Research Program of China (973 Program) under Grant no. 2014CB049500 and the Key Technologies R&D Program of Anhui Province under Grant no. 1301021005.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.ymssp.2015.06.004.

References

[1] B.C. Faulkner, F. Barton, T. Baber, W. Mckeel Jr., Determination of Bridge Response Using Acceleration Data, Technical Report, 1996.
[2] K.-T. Park, S.-H. Kim, H.-S. Park, K.-W. Lee, The determination of bridge displacement using measured acceleration, Eng. Struct. 27 (3) (2005) 371–378.
[3] Y.P. Zhang, H.N. Zhu, W.L. Zhou, H.F. Liu, Application of the Fourier transform in electronic speckle photography, Exp. Mech. 42 (1) (2002) 18–24.
[4] M. Meo, G. Zumpano, X.L. Meng, E. Cosser, G. Roberts, A. Dodson, Measurements of dynamic properties of a medium span suspension bridge by using the wavelet transforms, Mech. Syst. Signal Process. 20 (5) (2006) 1112–1133.
[5] L. Xu, J.J. Guo, J.J. Jiang, Time–frequency analysis of a suspension bridge based on gps, J. Sound Vib. 254 (1) (2002) 105–116.
[6] H.H. Nassif, M. Gindy, J. Davis, Comparison of laser doppler vibrometer with contact sensors for monitoring bridge deflection and vibration, Ndt E Int. 38 (3) (2005) 213–218.
[7] D.L.B.R. Jurjo, C. Magluta, N. Roitman, P.B. Goncalves, Experimental methodology for the dynamic analysis of slender structures based on digital image processing techniques, Mech. Syst. Signal Process. 24 (5) (2010) 1369–1382.
[8] T. Wu, Y. Peng, H. Wu, X. Zhang, J. Wang, Full-life dynamic identification of wear state based on on-line wear debris image features, Mech. Syst. Signal Process. 42 (1–2) (2014) 404–414.
[9] D. You, X. Gao, S. Katayama, Monitoring of high-power laser welding using high-speed photographing and image processing, Mech. Syst. Signal Process. 49 (1–2) (2014) 39–52.
[10] P.J. Figueroa, N.J. Leite, R.M.L. Barros, Tracking soccer players aiming their kinematical motion analysis, Comput. Vis. Image Underst. 101 (2) (2006) 122–135.
[11] F. Cheli, P. Mazzoleni, M. Pezzola, E. Ruspini, E. Zappa, Vision-based measuring system for rider's pose estimation during motorcycle riding, Mech. Syst. Signal Process. 38 (2) (2013) 399–410.
[12] F.C. Trigo, F.P.R. Martins, A.T. Fleury, H.C. Silva Jr., Identification of a scaled-model riser dynamics through a combined computer vision and adaptive kalman filter approach, Mech. Syst. Signal Process. 43 (12) (2014) 124–140.
[13] A.M. Wahbeh, J.P. Caffrey, S.F. Masri, A vision-based approach for the direct measurement of displacements in vibrating systems, Smart Mater. Struct. 12 (5) (2003) 785–794.

[14] J.J. Lee, M. Shinozuka, Real-time displacement measurement of a flexible bridge using digital image processing techniques, Exp. Mech. 46 (1) (2006) 105–114.
[15] C. Warren, P. Pingle, C. Niezrecki, P. Avitabile, Comparison of Image Based, Laser, and Accelerometer Measurements, Springer, New York, 2011, 15–21.
[16] Y. Fukuda, M.Q. Feng, Y. Narita, S. Kaneko, T. Tanaka, Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm, IEEE Sens. J. 13 (12) (2013) 4725–4732.
[17] E.S. Bell, J.T. Peddle, A. Goudreau, Bridge condition assessment using digital image correlation and structural modeling, Bridge Maint. Saf. Manag. Resil. Sustain. (2012) 330–337.
[18] W. Tong, Formulation of Lucas–Kanade digital image correlation algorithms for non-contact deformation measurements: a review, Strain 49 (4) (2013) 313–334.
[19] T. Siebert, R. Wood, K. Splitthof, High speed image correlation for vibration analysis, in: Proceedings of the Seventh International Conference on Modern Practice in Stress and Vibration Analysis, vol. 181, 2009.
[20] Y.F. Ji, C.C. Chang, Nontarget image-based technique for small cable vibration measurement, J. Bridge Eng. 13 (1) (2008) 34–42.
[21] S.W. Kim, N.S. Kim, Y.M. Kim, Application of vision-based monitoring system to stay cables, Bridge Maint. Saf. Manag. Resil. Sustain. (2012) 1116–1123.
[22] M.A. Iadicola, R.S. Zobel, J.M. Ocel, Quantitative evaluation of digital image correlation as applied to large-scale gusset plate experiments, Bridge Maint. Saf. Manag. Resil. Sustain. (2012) 1436–1443.
[23] B. Bascle, R. Deriche, Region tracking through image sequences, in: Proceedings of Fifth International Conference on Computer Vision, 1995, pp. 302–307.
[24] A.D. Jepson, D.J. Fleet, T.F. El-Maraghi, Robust online appearance models for visual tracking, IEEE Trans. Pattern Anal. Mach. Intell. 25 (10) (2003) 1296–1311.
[25] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: IJCAI, vol. 81, pp. 674–679.
[26] S. Baker, I. Matthews, Lucas–Kanade 20 years on: a unifying framework, Int. J. Comput. Vis. 56 (3) (2004) 221–255.
[27] S. Baker, I. Matthews, Equivalence and efficiency of image alignment algorithms, 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, Proceedings, 2001, pp. 1090–1097.
[28] G.D. Hager, P.N. Belhumeur, Efficient region tracking with parametric models of geometry and illumination, IEEE Trans. Pattern Anal. Mach. Intell. 20 (10) (1998) 1025–1039.
[29] M.J. Black, A.D. Jepson, Eigentracking: robust matching and tracking of articulated objects using a view-based representation, Int. J. Comput. Vis. 26 (1) (1998) 63–84.
[30] C.S. Fuh, P. Maragos, Motion displacement estimation using an affine model for image matching, Opt. Eng. 30 (7) (1991) 881–887.
[31] Y. Altunbasak, R.M. Mersereau, A.J. Patti, A fast parametric motion estimation algorithm with illumination and lens distortion correction, IEEE Trans. Image Process. 12 (4) (2003) 395–408.
[32] T.Y. Cheng, C. Herman, Involuntary motion tracking for medical dynamic infrared thermography using a template-based algorithm, in: Medical Imaging 2013: Image Processing, vol. 8669.
[33] T.Y. Cheng, C. Herman, Motion tracking in infrared imaging for quantitative medical diagnostic applications, Infrared Phys. Technol. 62 (2014) 70–80.
[34] A.J.H. Hii, C.E. Hann, J.G. Chase, E.E.W. Van Houten, Fast normalized cross correlation for motion tracking using basis functions, Comput. Methods Prog. Biomed. 82 (2) (2006) 144–156.