Smart Structures and Systems, *Vol. 31, No. 4 (2023) 421-436*
https://doi.org/10.12989/sss.2023.31.4.421

Target-free vision-based approach for vibration measurement and damage identification of truss bridges

Dong Tan [1a], Zhenghao Ding [2], Jun Li[*1] and Hong Hao [1,3]

[1] *Centre for Infrastructural Monitoring and Protection, School of Civil and Mechanical Engineering, Curtin University, Kent Street, Bentley, WA 6102, Australia*
[2] *Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China*
[3] *Earthquake Engineering Research and Test Center, Guangzhou University, Guangzhou, China*

(*Received August 26, 2022, Revised December 10, 2022, Accepted February 2, 2023*)

Abstract. This paper presents a vibration displacement measurement and damage identification method for a space truss structure from its vibration videos. Features from Accelerated Segment Test (FAST) algorithm is combined with adaptive threshold strategy to detect the feature points of high quality within the Region of Interest (ROI), around each node of the truss structure. Then these points are tracked by Kanade-Lucas-Tomasi (KLT) algorithm along the video frame sequences to obtain the vibration displacement time histories. For some cases with the image plane not parallel to the truss structural plane, the scale factors cannot be applied directly. Therefore, these videos are processed with homography transformation. After scale factor adaptation, tracking results are expressed in physical units and compared with ground truth data. The main operational frequencies and the corresponding mode shapes are identified by using Subspace Stochastic Identification (SSI) from the obtained vibration displacement responses and compared with ground truth data. Structural damages are quantified by elemental stiffness reductions. A Bayesian inference-based objective function is constructed based on natural frequencies to identify the damage by model updating. The Success-History based Adaptive Differential Evolution with Linear Population Size Reduction (L-SHADE) is applied to minimise the objective function by tuning the damage parameter of each element. The locations and severities of damage in each case are then identified. The accuracy and effectiveness are verified by comparison of the identified results with the ground truth data.

Keywords: damage identification; homography rectification; L-SHADE algorithm; sparse regularization; target-free feature detection; vibration displacement measurement

1. Introduction

Civil engineering structures are unavoidably subjected to loadings during their service. Some unexpected loadings, such as rough winds, earthquakes and other hazardous factors, may cause significant damage or even accidents. Therefore, structural damage identification, an active topic in Structural Health Monitoring (SHM), is essential to ensure structural safety in the civil engineering community.

With the development of data collection and measurement, methods based on structural vibration properties, such as natural frequencies (Lee 2009, Cawley and Adams 1979), mode shapes (Zhong and Oyadiji 2011), and acceleration responses (Lu and Law 2007), have been widely applied in structural damage identification due to their sensitivities to damage. However, these methods may have some limitations. For example, from the measurement aspect: i) The additional sensor weight of the contact-type sensors attached on structures may affect the vibration for mass-sensitive and light-weight structures; ii) The installation of sensors is limited by the locations of the target and site conditions of real-world structures; iii) For contactless sensors such as Laser Displacement Sensor (LDS), the labour-intensive installation and the requirement of fixed platforms also introduce some difficulties to the accurate acquisition of displacement response data; iv) The number of measurement points is limited, not only because of the expenses, but also the limitation of channels in data acquisition systems, therefore a full-field measurement is usually not realistic. Overall, many other factors limit the traditional sensors from realising the goal of structural monitoring and damage assessment by the existing theoretical methods.

Meanwhile, the recent development of computer science and vision technology has inspired the non-contact vibration displacement measurement for SHM (Hao *et al*. 2023). Computer vision-based measurement techniques can remedy the drawbacks in the mentioned traditional methods. Specifically, computer vision techniques serve as novel methods that can fulfill remote (Feng and Feng 2018, Tan *et al*. 2023) and full-field (Pan *et al*. 2009, Shao *et al*. 2021) vibration measurement simultaneously. Two major types of methods which are widely used in computer vision techniques, including template matching and feature tracking, have been developed widely and even adopted to

*Corresponding author, Associate Professor,
E-mail: junli@curtin.edu.au
[a] Ph.D. Candidate,
E-mail: Dong.Tan@student.curtin.edu.au

ISSN: 1738-1584 (Print), 1738-1991 (Online)

the field applications of SHM. For example, the template matching methods, such as Digital Image Correlation (DIC) (Reu *et al*. 2017, Chen *et al*. 2013, Helfrick *et al*. 2011), track the movement in terms of region. The displacement is calculated by matching the desired template region in the neighbour area among adjacent frames through optimising the image-related correlation functions. However, these methods often require some specific patterns, even preset artificial targets as template in some circumstances. These methods also require significant computation resources. On the other hand, the requirement of feature tracking is low, as long as there are stable key points on the tracking surface. Shi and Tomasi (1994) proposed an algorithm to identify key features in the first place. Thereafter, combined with KLT proposed by (Lucas and Kanade 1981, Tomasi and Kanade 1999) based on Optical Flow, algorithms such as Harris Corner (Harris and Stephens 1988, Yoon *et al*. 2016), Scale Invariant Feature Transform (SIFT) (Lowe 1999) and Speeded up robust features (SURF) (Bay *et al*. 2008, Kuddus *et al*. 2019) are developed as reliable methods for motion tracking and are broadly used in SHM. One prominent method among these algorithms is Features From Accelerated Segment Test (FAST) (Trajković and Hedley 1998). It has a significantly faster detection speed and better accuracy compared with the traditional DIC method, which makes it suitable for simultaneously tracking the objects in real-life measurement.

On the other hand, structural damage identification can be treated as an optimization problem by defining a structural parameter-related objective function. However, many objective functions are usually ill-posed and multi-dimensional, therefore the applications of traditional optimiasation methods (Lu and Wang 2017) may end up with a local minimum and cannot yield satisfactory identification results, which may lead to the failure of structural damage identification.

Differential Evolution (DE) (Storn and Price 1997, Storn 1996) is an evolutionary algorithm designated for optimisation. The control parameters of DE enable it to overcome complex problems (Das and Suganthan 2011), while these parameters are often problem dependent (Qin *et al*. 2009, Brest *et al*. 2006). Success-history based adaption is a mechanism by using the best historical memory settings. Theories showed that it guarantees a good performance of DE, but also ensures that the control parameters can be better chosen (Tanabe and Fukunaga 2013). The size of population plays an important role in computation convergence. Small population sizes tend to converge faster and better, but at the risk of reaching local optimum. Therefore, Linear Population Size Reproduction is introduced to balance the dilemma. The effectiveness of the newly built L-SHADE algorithm against others is evaluated by Tanabe and Fukunaga (2014).

In this study, a novel structural damage identification method based on the computer vision technique and L-SHADE algorithm is proposed. The synthetic videos of a truss structure under different structural conditions (Narazaki *et al*. 2020) are used for the vibration displacement measurement and damage identification. Homography rectification is applied to video frames, so that the inaccurate tracking caused from skewness of the camera shooting is overcome. Targets with better accuracy and stability for displacement tracking are extracted by using FAST feature detection implemented with adaptive threshold value strategy and tracked by KLT. SSI (Peeters and De Roeck 1999) is applied to analyse the obtained vibration displacement data and identify the natural frequencies and mode shapes of the truss structure. A multi-sample objective function is formulated based on Bayesian inference (Ding *et al*. 2019) using natural frequencies identified from the measured displacements. Then a sparse penalty component is also included in the objective function to handle the ill-posed situation and increase the speed of convergence. L-SHADE algorithm is used for the high-dimension optimisation of the initial numerical model updating and subsequent damage identification for truss under different damage states. The effectiveness of the proposed method is demonstrated by using videos of the truss structure under different damage scenarios, and the accuracy is verified by the ground truth of displacement responses and damage severities.

2. Motion tracking methodology

The proposed target-free vision-based approach for vibration displacement measurement and damage identification of spatial truss structures follow these basic steps: (i) Source videos pre-processing (for cases that image plane is not parallel to the truss front plane); (ii) Target-free vision-based motion tracking and displacement acquisition; (iii) Operational modal analysis from measured vibration data; (iv) Damage identification by model updating. The following contents in this section will cover the theoretical background and development of these procedures.

2.1 Homography rectification

Additional procedures must be undertaken for cases with the structural plane being not parallel to the image plane, since the scale factor is not a constant when the plane is skew in these cases. Besides, directional misalignment will occur when the horizontal and vertical axes are not overlapped.

A homography rectification (Loop and Zhang 1999) is adopted on the object plane to address this problem. As shown in Fig. 1, for two different image planes π and π', all co-existing points on the images can be linked by homography transformation **H**. For object points p on the

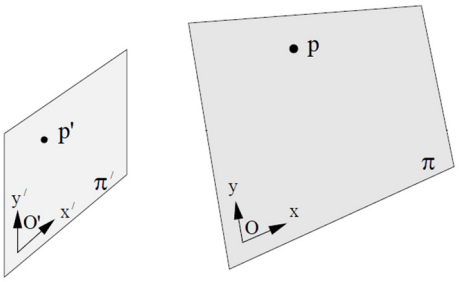


Fig. 1 Points mapping of homography

plane π, and p' on the plane π', the points pairs are mapped by p' = **H**p.

In rectification of the skew image, **H** is derived from the points pairs {p, p'} firstly, and is applied on target plane π. The matrix of homography transformation requires a minimal number of 4 pairs of point connections between the target image and reference image (Zhang 2003).

2.2 Target-free vision-based feature detection and tracking

The goal of vision-based vibration tracking is to capture the time-history displacement responses. The analysis is started with the first frame of the videos. By drawing some user-defined bounding boxes, the Region of Interest (ROI), namely the critical features of interest to be tracked throughout the entire video clip, are detected within the area by feature detection algorithms. In this research, the FAST algorithm (Trajković and Hedley 1998) is adopted, which has a better detection accuracy over the traditional Harris Corner detector (Harris and Stephens 1988), and the calculation speed is also guaranteed.

FAST corner detector applies on a circle of 16 pixels to determine whether a candidate point p is a corner point. Every pixel in the circle is labeled from integer 1 to 16 by clockwise, as shown in Fig. 2. If a set of N contiguous pixels in the circle are all brighter than the intensity of candidate pixel p (represented by I_p) over a threshold value of intensity t, or all darker than the intensity of candidate pixel p minus the threshold t, then point p is defined as a corner. The conditions can be written as

$$\forall x \in N,\ I_x > I_p + t \quad \text{or} \quad \forall x \in N,\ I_x < I_p - t \tag{1}$$

When either of the two conditions is met, candidate point p is classified as a corner. N is usually chosen as 12 (Rosten and Drummond 2005).

The FAST algorithm does not require comparing all pixels on the circumference with the one in the central place, thus many non-feature points can be quickly eliminated, and calculation time is reduced. However, the drawbacks are also obvious: i) for the images obtained under different circumstances, their contrast and noise situations exhibit differences. Therefore, their threshold values t should be different as well. If a fixed threshold is used, the robustness of the algorithm will be worse; b) due to the similarity between the feature points and their neighbourhoods, those neighbor points are also detected as feature points, resulting in 'feature blocks', which renders the high aggregation rate of feature points and reduces the performance of the algorithm.

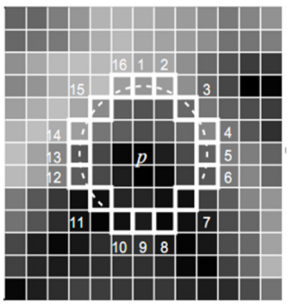


Fig. 2 The pixels used by the FAST corner detector

To deal with the problem, an adaptive threshold value strategy is adopted in the FAST algorithm. Assuming p_0 is the candidate point, then p_0 is treated as the center to formulate a square area. The adaptive local threshold value t_0 can be calculated as

$$t_0 = k \cdot \frac{\sum_{i=1}^{n} I_{max} - \sum_{i=1}^{n} I_{min}}{n\bar{I}} \tag{2}$$

where $I_{\max}$, $I_{\min}$ and $\bar{I}$ respectively denote the maximum, minimum and mean value of grey scale in the area, n is the number of pixels in the ROI, and k is a proportion ratio, which is generally taken as 3 (Rosten and Drummond 2005). The candidate feature points can thus be further judged by the unique threshold value inside each ROI. Consequently, the feature points selected are more stable, and the 'feature blocks' phenomena are reduced to some extent.

Given that the motion of structure is not very large, the optical flow-based tracking method can be applied on features tracking. Once the features are selected in the initial frame by FAST, the Kanade-Lucas-Tomasi (KLT) (Tomasi and Kanade 1999) method is used to track the features.

2.3 Vibration displacement response acquisition

According to the descriptions of the provided videos (Narazaki *et al*. 2020), all videos follow the standard pin-hole camera model, and are calibrated during synthesis of videos, hence no calibration for radial distortion or skew of lens is needed. The tracked displacements are converted directly from pixel into physical unit, i.e., meter, by multiplying the scale factor λ.

It is noted that the initial tracking results d_c only represent the position histories of the feature point, and they are expressed in pixel unit. Besides, every feature in the first frame is usually not on the origin of image plane coordinates, and its location is not (0, 0). Therefore, displacement should be offset by its initial position in the first frame of the video and then multiplied by the scale factor

$$d_w(i) = \lambda[d_c(i) - d_{ini}] \tag{3}$$

where d_{ini} is the coordinate of corresponding feature in the initial frame.

2.4 Scale factor adoption

Scale factor is the ratio that transfers the digital unit into physical real-world unit. Since the vertical displacement dominates the structural vibration motions in this study, the transverse vibration (parallel to the depth of field direction) is negligible compared with the vertical one. The façade of the spatial truss can be treated as a flat plane during vibration. Therefore, a constant scale factor is established when the principle axis is perpendicular to the object plane

(image plane and façade of truss are parallel). In reality, it is usually calculated by counting the number of pixel that covers a known physical distance as follows

$$\lambda = \frac{D}{N} \quad \left(\text{unit:}\frac{mm}{pixel}\right) \tag{4}$$

In this study, the scale factor of truss component in each video circumstance may vary. Each scale factor should be calculated based on the first frame of each case video, before it is adopted into displacement conversion. Cases like 'Case undamaged', 'Case 1' and 'Case 6' where the camera image plane is parallel to the truss front plane, can be calculated from original videos directly, other cases will be based on the images after homography rectification.

3. Damage identification method

3.1 Finite element damage model

Structural damage is represented by the reduction of the elemental stiffness in the truss components. The damage can be quantified by multiplying the elemental stiffness matrices with a ratio vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{N_{nel}}]$ in the finite element model of the truss. Therefore, the global stiffness matrix $\mathbf{K}$ can be formulated as

$$\boldsymbol{K}(\boldsymbol{\alpha}) = \sum_{i=1}^{N_{nel}} \alpha_i \boldsymbol{k}_i \tag{5}$$

where $\mathbf{k}_i$ is the stiffness matrix of the i^{th} element under the intact state, and N_{nel} is the total number of structural elements. For the ratio α_i, that corresponding element i is intact if $\alpha_i = 1$, and the element is totally damaged if $\alpha_i = 0$.

3.2 Model updating objective function

Structural damage results in the changes in structural vibration properties, such as natural frequencies and mode shapes, etc. Conversely, the alterations of natural frequencies can be used to construct the structural damage-related objective function, and damage identification is usually transferred to the minimisation of the corresponding objective function.

In practice, natural frequencies of structures are easily identified from the recorded structural responses, while the corresponding mode shapes are relatively difficult to acquire (Chen and Yu 2017). In this study, only 8 spans of truss are covered by the field of view of the video, therefore the operational mode shapes are not of the complete structure, only the frequencies are utilised to formulate the objective function in each case.

To improve the capacity against noise effect during optimisation for damage identification, the Bayesian inference item is introduced into the objective function. This offers a rigorous process for uncertainty quantification, the posterior density function (PDF) of the stiffness parameter vector $\boldsymbol{\alpha}$ in Eq. (5), can be obtained as

$$p(\boldsymbol{\alpha}|\boldsymbol{D}) = c \cdot p(\boldsymbol{D}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) \tag{6}$$

where $p(\boldsymbol{\alpha}|\mathbf{D})$ represents the PDF of $\boldsymbol{\alpha}$ given the vibration characteristics $\mathbf{D}$, $p(\mathbf{D}|\boldsymbol{\alpha})$ stands for the likelihood function based on $\boldsymbol{\alpha}$, and $p(\boldsymbol{\alpha})$ is the prior PDF of parameters $\boldsymbol{\alpha}$ based on observation or assumptions of model, c is a constant which makes the integral of PDF to be 1.

For structural damage identification in this study, vibration characteristics $\mathbf{D}$ include the identified natural frequencies $\boldsymbol{\omega}$ from the measured displacement responses. Furthermore, the prior distribution function $p(\boldsymbol{\alpha})$ of elemental stiffness parameters $\boldsymbol{\alpha}$ can be assumed as a uniform distribution. By using Bayesian inference, the natural frequencies are taken with N_s samples as $\boldsymbol{\omega} = [\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_{N_s}]$, and $\boldsymbol{\omega}_s = [\omega_{s,1}, \omega_{s,2}, \ldots \omega_{s,Nm}]$ is the vector of natural frequencies in the ascending order of the sth measurement. The likelihood functions of the frequency data are assumed to be independent, then the principle of maximum entropy is adopted, with Gaussian distributions applied on these data (Chen and Yu 2017). Therefore, the PDF of frequency parameter $\omega_{s,m}$ can be expressed as

$$p(\omega_{s,m}|\boldsymbol{\alpha}) = c_1 \cdot exp\left[-\frac{(\omega_{s,m} - \omega_m^A)^2}{2\sigma_m^2}\right] \tag{7}$$

where ω_m^A represents the m-th order of structural natural frequency, σ_m is the standard deviation of the m-th frequency in the N_s sample, c_1 is another constant that makes sure that the integral of Eq. (7) is 1.

The obtained natural frequencies are assumed to be independent, therefore the likelihood function Eq. (6) by combining Eq. (7) is

$$\begin{aligned} p(\boldsymbol{\alpha}|\boldsymbol{D}) &= c \cdot p(\boldsymbol{\alpha}) \cdot \prod_{s=1}^{N_F} \prod_{m=1}^{N_m} p(\omega_{s,m}|\boldsymbol{\alpha}) \\ &= c \cdot p(\boldsymbol{\alpha}) \cdot exp\left[-\sum_{s=1}^{N_F} \sum_{m=1}^{N_m} \frac{(\omega_{s,m} - \omega_m^A)^2}{2\sigma_m^2}\right] \end{aligned} \tag{8}$$

where N_m is the total number of used natural frequencies.

Eq. (8) shows the natural frequency based Bayesian conditional PDF. The goal is to maximise this equation based on the test data. With the prior PDF $p(\boldsymbol{\alpha})$ of $\boldsymbol{\alpha}$ considered having a uniform distribution, the maximisation is converted into the minimisation of the exponent part of Eq. (8) (without the minus sign).

To overcome the difficulty with a limited number of measured natural frequencies and improve the structural damage identification performance, a sparse penalty function is added to cope with the ill-posed situation. Tikhonov regularisation item (Tikhonov 1963) could be introduced to increase the identification accuracy (Titurus and Friswell 2008, Zhang and Xu 2016). However, the Tikhonov regularisation may induce excessive smooth results. Recently, Xu *et al*. (2010) found out that the $L_{0.5}$-norm has a better sparsity and robustness than the traditional L_1-norm. Inspired by this, the identification of stiffness parameters vector $\boldsymbol{\alpha}$ is realised by minimising a combination of Bayesian conditional probability function

and an additional $L_{0.5}$-norm-based sparsity regularization item $\|1\text{-}\boldsymbol{\alpha}\|$, based on intrinsic frequencies, which is given as

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{argmin}\left\{\sum_{s=1}^{N_s}\sum_{m=1}^{N_m}\left[\frac{(\omega_{s,m}-\omega_m^A)^2}{2\sigma_m^2}\right] + \tau\|1-\boldsymbol{\alpha}\|_{0.5}\right\} \quad (9)$$

where $\omega_{s,m}$ denotes the m^{th} natural frequency in the numerical finite element model analysis, τ is a factor that controls the convergence properties of sparse regularisation. A small τ will pose a higher penalty and lead to an over-fitting solution. Conversely, a larger τ will result in losing data fidelity. The discrepancy principle (DP) rule (Hou *et al*. 2018) is employed here to choose the optimal regularization parameter τ.

The essence of the first term of objective function is the Bayesian inference (Ding *et al*. 2019), which provides a rigorous process for uncertainty quantification, and natural frequencies from multiple samples are adopted to make full use of the data sets, not only the average value (Chen and Yu 2017).

3.3 Success-history based Adaptive Differential Evolution with linear population size reduction (*L-SHADE*)

Normal optimization techniques could obtain some local minimals, which may lead to poor identification results. The high-performance algorithm Differential Evolution (DE) (Storn and Price 1997, Das and Suganthan 2011) is a competitive evolution algorithm in optimization problems. DE is widely known since its control parameters can solve complex problem, which is also the drawback, because these parameters are usually problem- dependent (Storn and Price 1997, Das and Suganthan 2011). Success-history based adaptation is a mechanism of parameter tuning (Tanabe and Fukunaga 2014). It uses good historical memory to generate new parameters, and it is also applicable to DE and leads to the inspiration of L-SHADE. The details of L-SHADE algorithm are provided below.

3.3.1 Control parameters assignment from historical memory

Feasible solutions are represented as a set of vectors $\boldsymbol{x}_i = (x_1, \ldots, x_D)$, $I = 1, 2, \ldots, N$, where D is the dimension of the solution and N is the population size.

The Success-history based Adaptive Differential Evolution (SHADE) (Tanabe and Fukunaga 2013) provides the historical memory with H entries for parameters CR, F, M_{CR} and M_F in DE. Among them, $F \in [0,1]$ is a scaling factor that controls the magnitude of the differential mutation operator. $CR \in [0,1]$ is the crossover rate. M_{CR} and M_F are the memories of CR and F, which will update with generation as the optimisation proceed, and are set to 0.5 at the beginning of the optimization. In the updating process, the elements of CR and F for each x_i component are generated by index r_i selected from $[1, H]$ and can be expressed as

$$CR_i = \begin{cases} 0 & \text{if } M_{CR,r_i} = \perp \\ \text{randn}_i(M_{CR,r_i}, 0.1) & \text{otherwise} \end{cases} \quad (10)$$

$$F_i = \text{randc}_i(M_{F,r_i}, 0.1) \quad (11)$$

If CR is outside [0, 1], it will be replaced by the nearest limit value (0 or 1). F_i is truncated to 1 when it is over 1, and will be repeatedly applied by Eq. (11) if it is less than 0 (Zhang and Sanderson 2009). If the corresponding history memory M_{CR,r_i} is empty (shown as $\perp$ in Eq. (10), the corresponding CR_i is set to 0.

3.3.2 Reproduction of trail vector

Afterwards, a mutant parameter $\boldsymbol{v}_{i,G}$ is generated from the population members by applying the current-to-*p*best/1 strategy, which is a variant of the current-to-best/1 strategy. In the current-to-best/1 strategy, only the historical individual with best fitness is used to produce the new individual. While for the current-to-*p*best/1 strategy, selected element for calculation in the current generation is chosen from p elements with top fitness values. In I-SHADE, the current-to-*p*best/1 strategy can be expressed as

$$\boldsymbol{v}_{i,G} = \boldsymbol{x}_{i,G} + F_i \cdot (\boldsymbol{x}_{pbest,G} - \boldsymbol{x}_{i,G}) + F_i \cdot (\boldsymbol{x}_{r_1,G} - \boldsymbol{x}_{r_2,G}) \quad (12)$$

In this equation, the current-to-*p*best/1 strategy enables $\boldsymbol{x}_{pbest,G}$ to be selected from the top $N\times p$ ($p \in [0,1]$) members in the G^{th} generation randomly. The parameter p controls the extent of exploitation and exploration. r_1 and r_2 are random numbers from [1, N]. When vector element $v_{j,i,G}$ exceeds the search boundaries $\left[x_j^{min}, x_j^{max}\right]$, then the below correction is performed (Zhang and Sanderson 2009)

$$v_{j,i,G} = \begin{cases} (x_j^{min} + x_{j,i,G})/2 & \text{if } v_{j,i,G} < x_j^{min} \\ (x_j^{max} + +x_{j,i,G})/2 & \text{if } v_{j,i,G} > x_j^{max} \end{cases} \quad (13)$$

After the mutant vector $\boldsymbol{v}_{i,G}$ is calculated in Eq. (12), it will be crossed with the parent $x_{i,G}$ by Binomial Crossover to generate the trial vector $\boldsymbol{u}_{i,G}$ as

$$u_{j,i,G} = \begin{cases} v_{j,i,G} & \text{if rand}[0,1) \le CR_i \text{ or } j = j_{rand} \\ x_{j,i,G} & \text{otherwise} \end{cases} \quad (14)$$

where j_{rand} is a decision variable index selected randomly from $[1, D]$.

3.3.3 Choice for new generation

After all the trail vectors $\boldsymbol{u}_{i,G}$ are calculated, the objective function will compare each parent vector $\boldsymbol{x}_{i,G}$ against the corresponding trail vector $\boldsymbol{u}_{i,G}$ and the new generation is obtained judging by the fitness values as

$$x_{i,G+1} = \begin{cases} u_{i,G} & \text{if } f(\text{u}_{i,G}) \le f(x_{i,G}) \\ x_{i,G} & \text{otherwise} \end{cases} \quad (15)$$

3.3.4 External archive

An optional external archive $\boldsymbol{A}$ is used for maintaining diversity (Zhang and Sanderson 2009) so that parents $\boldsymbol{x}_{i,G}$ which are worse than the trail vectors $\boldsymbol{u}_{i,G}$ are preserved. In

that case, $\boldsymbol{x}_{i,G}$ rejected by Eq. (15) is collected into $\boldsymbol{A}$, $x_{r_2,G}$ in Eq. (12) is selected by $\boldsymbol{P} \cup \boldsymbol{A}$, the union of population $\boldsymbol{P}$ and archive $\boldsymbol{A}$.

3.3.5 Historical-memory update

CR_i and F_i that manage to generate a better trail vector are recorded as S_{CR} and S_F in each generation. The memory content M_{CR} and M_F for each element k in the G+1 generation are updated by the following algorithm:

Algorithm 1: Memory update algorithm

If $S_{CR} \neq 0$ **and** $S_F \neq 0$ **then**
1. **If** $M_{CR,k,G} = \perp$ **or** $\max(S_{CR}) = 0$ **then**
2. $M_{CR,k,G+1} = \perp$;
3. **else**
4. $M_{CR,k,G+1} = \text{mean}_{WL}(S_{CR})$; **End If**
5. $M_{F,k,G+1} = \text{mean}_{WL}(S_F)$;
6. k++
7. If $k > H$, $k = 1$;
8. **else**
9. $M_{CR,k,G+1} = M_{CR,k,G}$;
10. $M_{F,k,G+1} = M_{F,k,G}$;

End If

The index k determines the position of update, its value is initialized to 1 and incremented when new elements are inserted into histories. If $k > H$, it is set to 1 again. The weighted Lehmer mean (Peng *et al*. 2009) is computed as

$$\text{mean}_{WL}(S) = \frac{\sum_{k=1}^{|S|} \omega_k \cdot S_k^2}{\sum_{k=1}^{|S|} \omega_k \cdot S_k} \tag{16}$$

$$\omega_k = \frac{\Delta f_k}{\sum_{l=1}^{|S_{CR}|} \Delta f_l} \tag{17}$$

where Δf_k is the amount of fitness (value of objective function) improvement of the k_{th} element.

3.3.6 Linear Population Size Reproduction (LPSR)

Population Size Reduction is effective in improving performance of evolutional algorithm (Brest and Maučec 2008, Vlašić *et al*. (2019), thus it is also applicable to SHADE. Here, after the G^{th} generation, LPSR reduces the population to match a linear function as below

$$N_{G+1} = round\left[\left(\frac{N^{min} - N^{init}}{MaxNFE}\right) \cdot NFE + N^{init}\right] \tag{18}$$

where N^{min} and N^{init} are the smallest population value at the final stage and its initial value, respectively; *NFE* is the number of fitness evaluations in the current generation, and *MaxNFE* is its maximum value. Therefore, the ($N_G - N_{G+1}$) worst individuals will be deleted from population at the G^{th} generation when $N_G > N_{G+1}$.

This population reduction strategy reduces the computational demand of SHADE and makes it converge fast. Therefore, the overall steps and strategies of L-SHADE are described below.

Algorithm 2: L-SHADE algorithm

Initialisation
1. $G = 1$, $N_G = N^{\text{init}}$, Archive $\boldsymbol{A} = \varnothing$;
2. Initialize population $\boldsymbol{P}_G = (\boldsymbol{x}_{i,G}, ..., \boldsymbol{x}_{N,G})$ randomly;
3. Set M_{CR} and M_F to 0.5;

Main loop
4. **While** The termination criteria are not met **do**;
5. S_{CR} and $S_F = \varnothing$;
6. **for** $i = 1$ to N **do**
7. r_i = rand[1, H];
8. **If** $M_{CR,r_i} = \perp$, $CR_{i,G} = 0$, else $CR_{i,G}$ = rand n_i (M_{CR,r_i}, 0.1);
9. $F_{i,G}$ = rand c_i (M_{F,r_i}, 0.1);
10. Generate trail vector $\boldsymbol{u}_{i,G}$ according to current-to-*p*best/1 strategy
11. **for** $i = 1$ to N **do**
12. **If** $f(\boldsymbol{u}_{i,G}) \leq f(\boldsymbol{x}_{i,G})$ **then**
13. $\boldsymbol{x}_{i,G+1} = \boldsymbol{u}_{i,G}$; **else**
14. $\boldsymbol{x}_{i,G+1} = \boldsymbol{x}_{i,G}$; **End if**
15. **If** $f(\boldsymbol{u}_{i,G}) < f(\boldsymbol{x}_{i,G})$ **then**
16. $\boldsymbol{x}_{i,G} \rightarrow \boldsymbol{A}$;
17. $CR_{i,G} \rightarrow S_{CR}$, $F_{i,G} \rightarrow S_F$; **End if**
18. If necessary, delete randomly selected individuals from the archive $\boldsymbol{A}$, such that the archive size is $|\boldsymbol{A}|$.
19. Update M_{CR} and M_F according to Algorithm 1;

Optional LPSR strategy
20. Calculate N_{G+1} according to Eq. (18)
21. **If** $f(\boldsymbol{u}_{i,G}) < f(\boldsymbol{x}_{i,G})$ **then**
22. Sort individuals in P based on their fitness and delete the lowest ($N_G - N_{G+1}$) members, resize archive size $|\boldsymbol{A}|$ according to new $|\boldsymbol{P}|$;
23. Proceed with next generation G+1;

3.4 Workflow of the proposed method

The general steps of the proposed damage identification method based on target-free vibration displacement measurements can be summarized as following steps and illustrated by Fig. 3,

- Homography rectification for videos of special cases by using the first frames of target video and reference video;
- Scale factor for each case is calculated based on the dimension of an object in the first frame and its physical dimension;
- The ROIs are chosen in every video around each area, afterwards, the target points for tracking are selected by performing FAST algorithm with adaptive threshold inside the ROI;
- The KLT algorithm tracks the positions of selected features, with output in the unit of pixel, and subsequently converted into physical unit by scale factor;
- Operational frequencies and mode shapes are identified from obtained nodal displacements, by using Stochastic Subspace Identification (SSI);
- Finite element model is built and analytical model updating is conducted from structural properties under undamaged state;

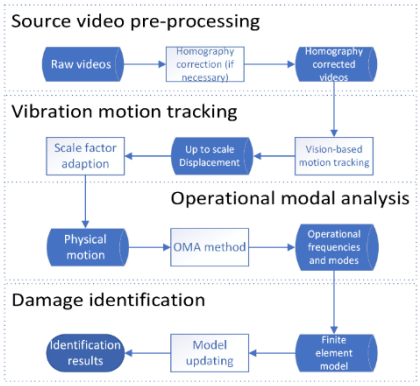


Fig. 3 Workflow of the proposed method

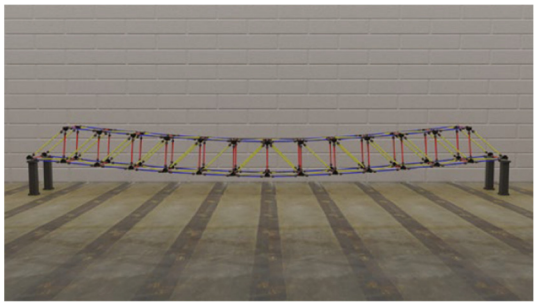
- Construction of the objective function by identified natural frequencies, and finite model updating is conducted by L-SHADE;
- Identification results of damage severity are obtained from the updated damage indices.

4. Method validation by the synthetic truss dataset

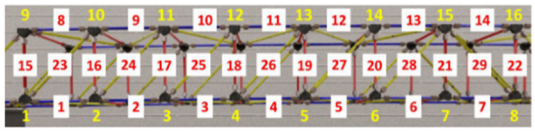
In this section, the accuracy and performance of the proposed approach for vibration displacement measurement and damage identification are validated. The tracking results of vibration displacement responses from the vision-based method are compared with the ground truth data, and errors are presented by several criteria. Operational modal analysis is conducted to obtain the operational natural frequencies and mode shapes of the truss structure. Identification of damage location and severity are then carried out by the identified vibration characteristics using optimisation algorithms. The following sections will present the experimental model and set-up, and then present and discuss the obtained displacement measurement and damage identification results.

4.1 Setup of simulation

The spatial truss structure for experimental verifications has 56 joints and 160 elements in total. The lower deck of the bridge has 14 spans and the upper deck has 12 spans, as shown in Fig. 4(a). While the field of view in every video covers only the 2nd to 9th spans of the bridge, including 16 nodes and 29 elements on the front façade, these spans are treated as the analytical targets in this study. The numbers of nodes are marked in yellow color, and the elements in red color, as shown in Fig. 4(b). In other words, their corresponding elemental stiffness matrices are parameterised and multiplied with damage indices for identification following the sequence of marked numbers in Fig. 4(b). The longitudinal and transverse elements have a length of 0.3937 m, and vertical elements are 0.4 m long. The truss structure is simply-supported with the left side pinned on the ground, and the right side placed on a roller. The cross section of the truss is a hollow circular section, which is identical for all elements, with an outer diameter of 0.01554 m and an inner diameter of 0.01087 m. The



(a) Diagram of the whole structure



(b) Available part in the videos

Fig. 4 Physical model of the spatial truss

Young's modulus, shear modulus and mass density are approximately 18.5 GPa, 79.3 GPa and 8000 kg/m^3, respectively.

The duration of each video is 240 s, and the frame rate of videos is 120 fps with a resolution of 1920 × 1080 pixels. Videos after homography rectification may exceed this resolution. The Nyquist's limit of frequency is 120 fps/2 = 60 Hz, therefore the vibration characteristics within 60 Hz can be obtained. In all the damage cases, the structure is excited by ambient forces, where vertical-direction displacement response dominates the vibration. The results of the obtained vibration displacement responses and mode shapes, which refer to the partial vibration data in the vertical direction, are presented in this study.

Apart from the undamaged state, there are 10 different damage scenarios, which is marked 'Case X' with 'X' denotes the order of damage scenario. The truss structure in all cases has only one damage on one of the marked components in Fig. 4(b), and the reduction of damage is over 40%.

4.2 Homography rectification for unparalleled cases

For some special cases including Case 3, Case 4, Case 5, Case 8, Case 9 and Case 10, the front façade plane of the truss is not parallel to the image plane. Hence not like other cases, videos of these cases must be processed by performing a homography transformation before conducting the tracking for obtaining vibration displacement responses.

To illustrate its effect, the first frame of the video clip of 'Case 10' is chosen as an example for performing homography rectification. The first frame of the video for 'Case Undamaged' which has a parallel image plane, is chosen as the reference image. During points matching, point correspondences should be the same physical points on structures, which means that the target image and reference image where these points located on, should be in the same plane in real-world. In this study, the motion measurement of joints is essential, therefore corner points on each node joint are a preferable choice to assure better

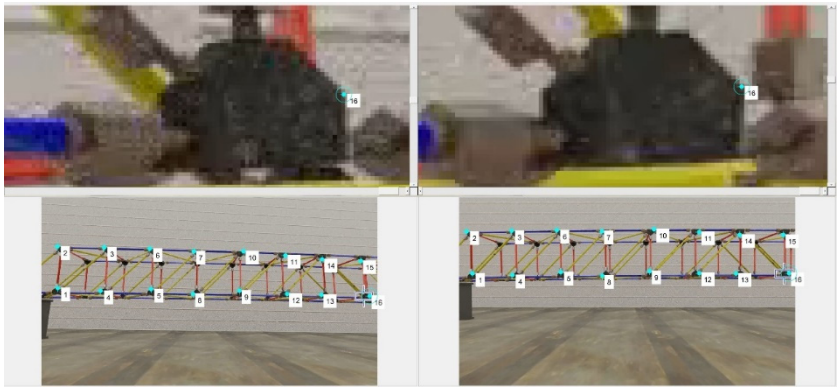
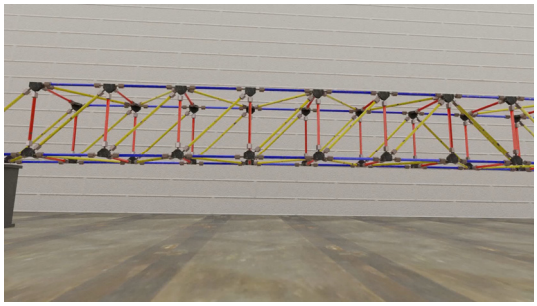
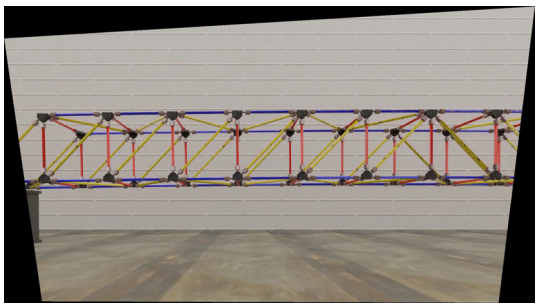


Fig. 5 Points correspondence selection



(a) Before Processing



(b) After homography rectification

Fig. 6 Comparison of frames from Case 10 with homography rectification

transformation and tracking results. To achieve a better accuracy, 16 points pairs around the corner of all joints are chosen to optimise the homoraphy matrix, as shown in Fig. 5. After homography transformation, it can be observed that the original skew quadrilateral of the truss of 'Case 10' as shown in Fig. 6(a), becomes as rectangular as the reference image, which is similar as 'Case Undamaged' as shown in Fig. 6(b). The shooting stance of target frame after processing is deemed to be the same as the reference frame, the scale factor on the target plane is uniform hereafter.

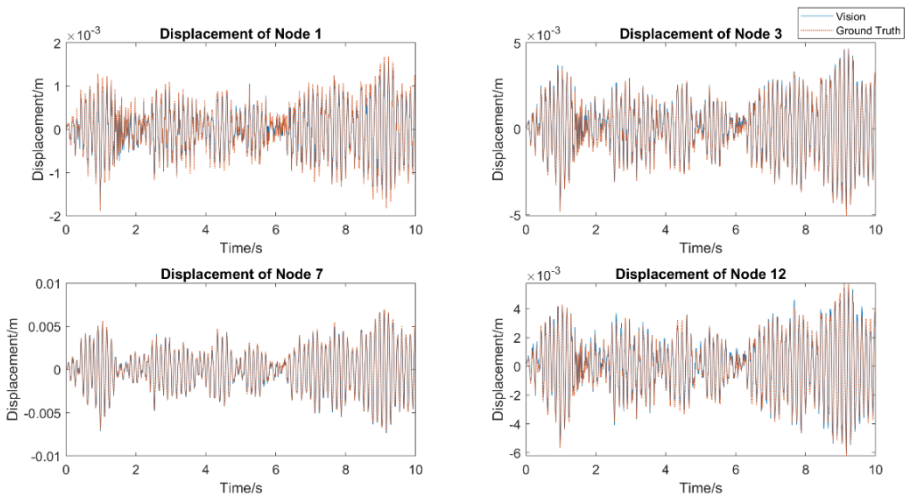
It should be noted that the above properties hold only for a plane, as stated in Section 2.1. Therefore, all the points correspondences cannot be selected outside the façade plane of the truss. The black edges on the sides of the processed image, as shown in Fig. 6(b), are not caused from homography transformation. It is because that the digital image file must be stored as a rectangle, the operating system will automatically fill and present the void pixel area in black colour, which ensures that the rectified graphic is covered by the whole rectangle image. The image size (pixel number) of each case must be counted again and adjusted, since image size is enlarged by these black edges.

Since the camera is stationary, the homography rectification is applied on every frame of the video. Afterwards, a new video is created by combining these processed frames sequence by sequence. The displacement measurements of these cases are achieved by detecting and tracking the key points in these new videos. The effect of homography rectification as shown in Fig. 6(b), can be observed from the provided rectified sample video of damage Case 9 and Case 10.

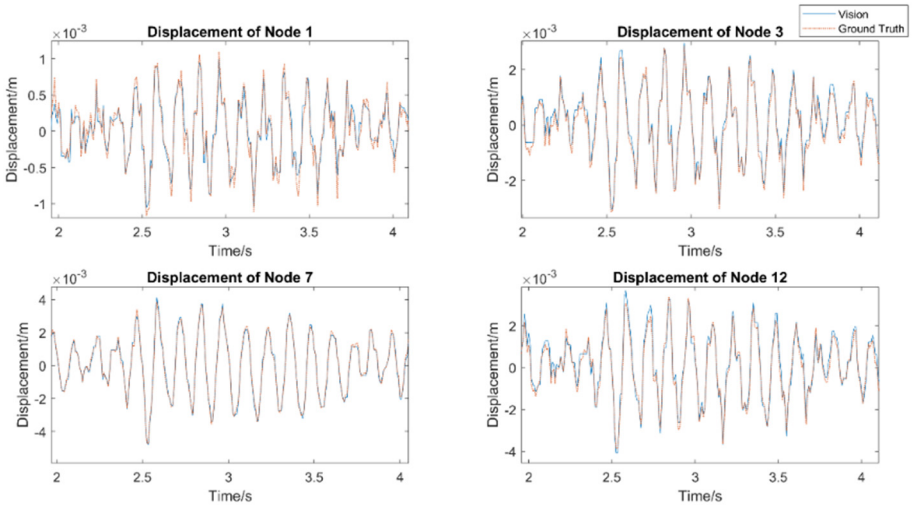
4.3 Vibration displacement measurement

To demonstrate the effectiveness of the proposed method on displacement responses measurement and modal identification, the results of 'Case 1' are presented in Sections 4.3 and 4.4. Four key points are chosen to demonstrate the effectiveness of the proposed method, namely nodes 1, 3, 7 on the lower deck and 12 on the upper deck. The vibration displacement responses are obtained by using the presented vision-based method in Section 2. Comparisons are made with ground truth data of the first 10 s of the video, as shown in Fig. 7(a). It can be observed that vision-based tracked displacement responses have a good match with ground truth data in the whole time duration, both on trends and amplitudes. A time interval between 2 s and 4 s is zoomed to present the details as shown in Fig. 7(b). Two vibration displacement responses from the proposed approach and ground truth correlate very well even in a short time duration for all nodes. There is no mismatch on the oscillation period when each node vibrates around the equilibrium position, even at small amplitudes around 6 s. This shows the accuracy and effectiveness of the proposed method on displacement response tracking.

To quantify the accuracy of the proposed tracking method, error analysis is conducted on these data of damage 'Case 1'. Apart from conventional error estimation criteria including correlation, standard deviation, three extra indices, Root Mean Square Error (RMSE), Coefficient of Determination (R^2) and Time Response Assurance Criterion (TRAC) (Baqersad *et al. 2015*) *are* used as follows



(a) Time histories over 10 seconds



(b) A zoom window of the duration between 2 s and 4 s

Fig. 7 Displacement responses comparison between the proposed method and ground truth data on nodes 1, 3, 7 and 12

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(d_v - d_g)^2}{N}} \tag{19}$$

$$R^2 = 1 - \frac{\sum(d_v - d_g)^2}{\sum(d_v - \mu_g)^2} \tag{20}$$

$$\text{TRAC} = \frac{(d_v \cdot d_g)^2}{{d_v}^2 \cdot {d_g}^2} \tag{21}$$

where d_v and d_g are the displacements obtained by the proposed vision-based method and ground truth, respectively, μ_g is the mean of the corresponding data series.

TRAC is a quantity to measure the correlation of two time domain signals. Therefore, when the target displacement response has a great correlation with the reference data, the TRAC value is high and close to 1, although the amplitudes of two groups of data may vary. The error analysis is conducted on all 16 nodal points with respect to the reference data, the results of five criteria can be found in Table 1 and Fig. 8. A great accuracy is observed between data of vision-based tracking method and ground truth. The RMSE values and standard deviations are largely below 3×10^{-4} m. R^2, correlation and TRAC values are over 0.96 except node 9. Results from the traditional Harris Corner are also conducted and analysed, as shown in Table 2 and Fig. 8. For most of nodes, the tracking results by adaptive threshold based FAST is better than the traditional Harris Corner, since smaller RMSE and Standard Deviation values, and greater R2, Correlation and TRAC values can be observed from the results of FAST with adaptive threshold value strategy than Harris Corner, as shown in Table 1 and Fig. 8.

Two tendencies are concluded: i) tracking results of points near the left fixed end are prone to be poorer. This is reflected by results of nodes 1 and 9, which have the worst results on the lower and upper deck respectively. Relatively better results are found on points located right near the middle span, no matter from amplitudes or details. This is because the lower bound of physical tracking accuracy lays on the pixel-wise resolution, namely scale factor in each case. With nodes located away from the fixed end and near

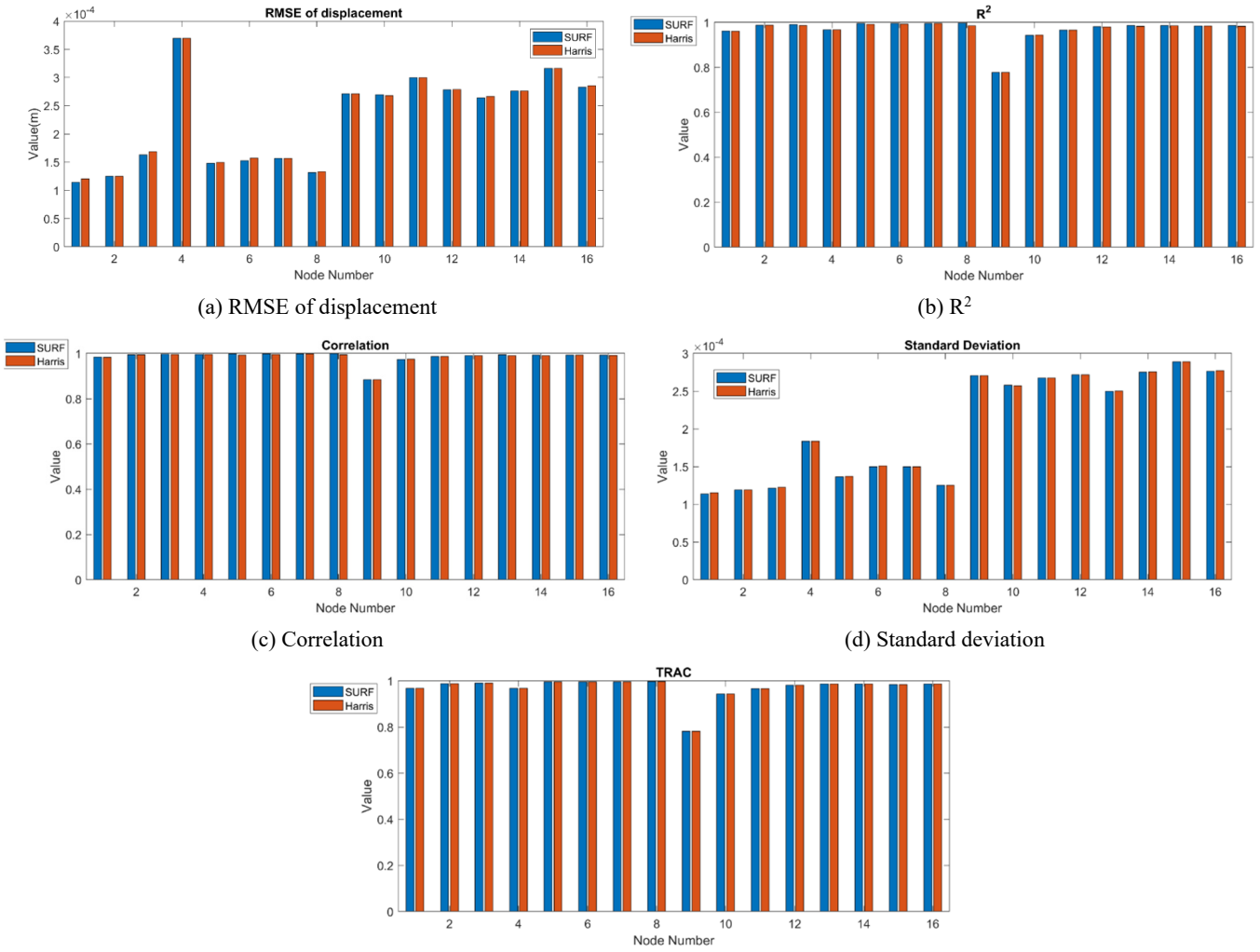


Fig. 8 Numerical error analysis of measured displacements of each node with reference data

Table 1 Error analysis for damage Case 1 by FAST

Node	RMSE (m)	R^2	Correlation	STD (m)	TRAC
1	1.1433×10^{-4}	0.9622	0.9839	1.1383×10^{-4}	0.9677
2	1.2492×10^{-4}	0.9879	0.9947	1.1922×10^{-4}	0.9882
3	1.6269×10^{-4}	0.9899	0.9972	1.2122×10^{-4}	0.9899
4	3.6935×10^{-4}	0.9672	0.9959	1.8372×10^{-4}	0.9679
5	1.4786×10^{-4}	0.9957	0.9982	1.3623×10^{-4}	0.9959
6	1.5258×10^{-4}	0.9960	0.9981	1.4963×10^{-4}	0.9960
7	1.5676×10^{-4}	0.9960	0.9982	1.4986×10^{-4}	0.9960
8	1.3150×10^{-4}	0.9970	0.9988	1.2523×10^{-4}	0.9973
9	2.7087×10^{-4}	0.7768	0.8844	2.7049×10^{-4}	0.7815
10	2.6952×10^{-4}	0.9426	0.9735	2.5779×10^{-4}	0.9426
11	2.9987×10^{-4}	0.9655	0.9870	2.6745×10^{-4}	0.9666
12	2.7844×10^{-4}	0.9808	0.9909	2.7159×10^{-4}	0.9810
13	2.6409×10^{-4}	0.9864	0.9941	2.4974×10^{-4}	0.9867
14	2.7594×10^{-4}	0.9870	0.9935	2.7517×10^{-4}	0.9870
15	3.1603×10^{-4}	0.9836	0.9934	2.8897×10^{-4}	0.9840
16	2.8291×10^{-4}	0.9860	0.9934	2.7625×10^{-4}	0.9861

the middle span, the vibration amplitudes obviously become larger and cover more pixels during vibration. Therefore, the deviations account for less part of the amplitudes, displacement time histories from vision-based method match better with ground truth data on numerical error analysis in these nodes. This phenomenon occurs every case. ii) The results from the lower deck are generally better than those from the upper deck. This is because the image quality of detected feature points is more stable for nodes on the lower deck. Overall, it can be concluded that the measured displacement responses have a good agreement with the ground truth, which indicates a good accuracy of the presented vision-based method for displacement measurement.

4.4 Operational modal identification

Operational modal identification is conducted to obtain natural frequencies and mode shapes of the structure. The frequency identification range is 0 to 60Hz with camera frame rate of 120 fps, the power spectrum density (PSD) and the stabilisation diagram obtained by SSI (Peeters and De Roeck 1999) by using vision tracking data and ground

Table 2 Error analysis for damage Case 1 by Harris Corner

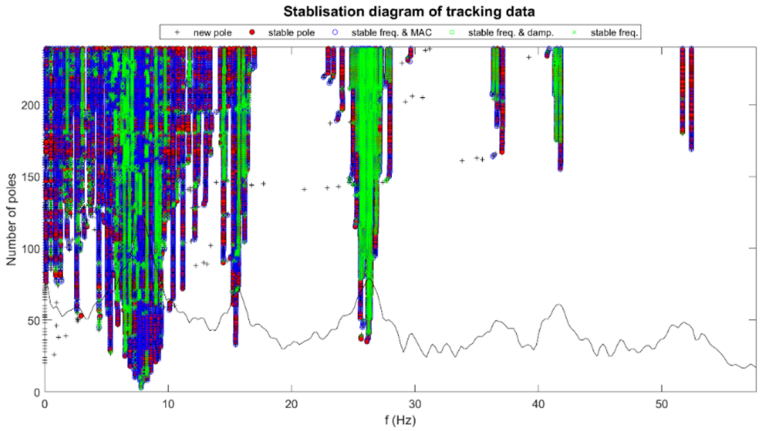
Node	RMSE (m)	R^2	Correlation	STD (m)	TRAC
1	1.2041×10^{-4}	0.9608	0.9826	1.1505	0.9520
2	1.2492×10^{-4}	0.9879	0.9947	1.1922	0.9882
3	1.6843×10^{-4}	0.9863	0.9960	1.2240	0.9886
4	3.6935×10^{-4}	0.9672	0.9959	1.8372	0.9679
5	1.4930×10^{-4}	0.9921	0.9930	1.3697	0.9868
6	1.5724×10^{-4}	0.9923	0.9950	1.5081	0.9909
7	1.5676×10^{-4}	0.9960	0.9982	1.4986	0.9960
8	1.3310×10^{-4}	0.9847	0.9945	1.2537	0.9949
9	2.7087×10^{-4}	0.7768	0.8844	2.7049	0.7815
10	2.6790×10^{-4}	0.9442	0.9748	2.5730	0.9439
11	2.9987×10^{-4}	0.9655	0.9870	2.6745	0.9666
12	2.7847×10^{-4}	0.9798	0.9901	2.7165	0.9807
13	2.6646×10^{-4}	0.9830	0.9898	2.5033	0.9805
14	2.7636×10^{-4}	0.9852	0.9904	2.7553	0.9835
15	3.1603×10^{-4}	0.9836	0.9934	2.8897	0.9840
16	2.8549×10^{-4}	0.9830	0.9912	2.7728	0.9830

truth data are compared and shown in Fig. 9. The overall trends of these two diagrams are similar. Combined with the PSD and the poles of the stabilisation diagram, six natural frequencies are identified from the peaks and listed in Table 3. These identified frequencies results are very accurate as they are all identical with the natural frequencies identified from ground truth.

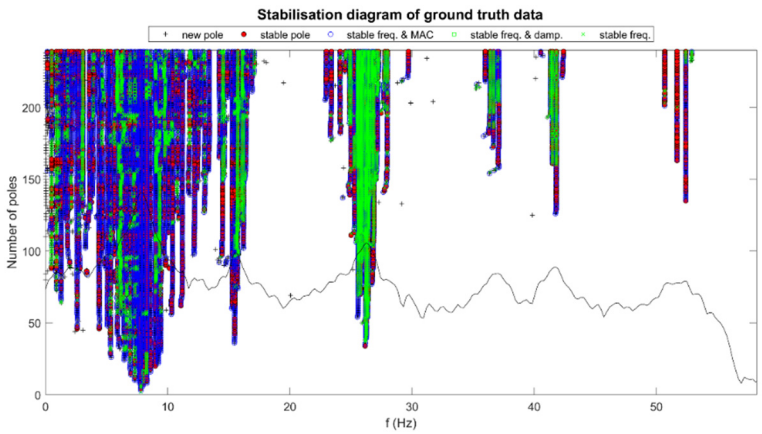
Table 3 Natural frequencies from displacement by the vision-based method and the relative errors

Mode order	1	2	3	4	5	6
Frequency (Hz)	7.73	15.47	26.25	37.03	41.72	52.41
Relative error (%)	0	0	0	0	0	0
MAC	0.9996	0.9505	0.9977	0.9836	0.9545	0.9903

The corresponding mode shapes for these six natural frequencies are shown in Fig. 10. It should be noted that these six mode shapes are only the modes that are excited in the vertical direction. It should be noted that these six operational frequencies are not the natural frequencies of the whole structure that are ranked in sequence. These are the frequencies that are excited in the vertical direction. Besides, only the partial mode shapes that contain the values of nodes from 1st to 8th spans, which corresponds to the available regions of the videos, are further presented and compared. All values are normalized regarding the available nodes. It can be seen from each sub-figure that mode shapes from two data groups have a very similar pattern on each



(a) Displacement data from vision-based method



(b) Ground truth data

Fig. 9 Stabilisation diagram and power spectrum density diagram

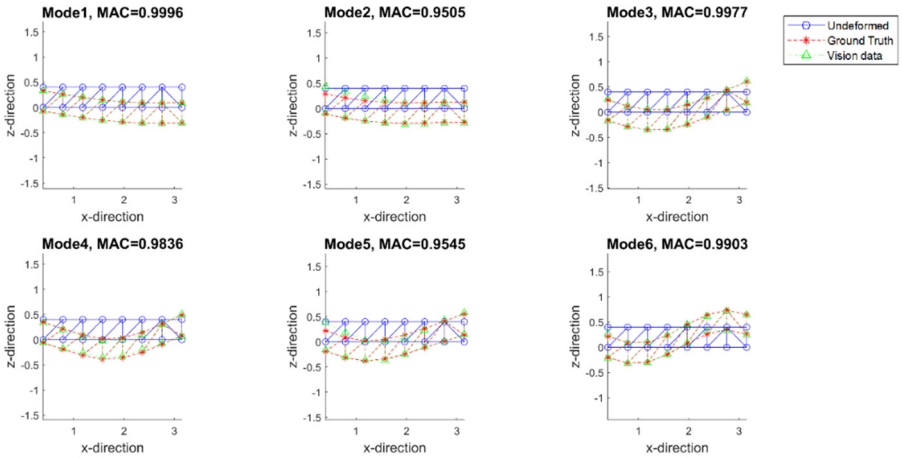


Fig. 10 First six identified operational mode shapes from ground truth and vision-based displacement data

Table 4 Comparison between natural frequencies from measured responses and updated finite element model

Mode order	Measured (Hz)	Updated (Hz)	Relative error
1	7.96875	7.9857	0.2127%
2	15.4688	15.4862	0.1125%
3	26.7188	26.7095	0.0348%
4	37.2656	37.3806	0.3086%
5	41.7188	41.7883	0.1666%
6	52.4087	52.4126	0.0074%

order and close magnitudes on each node. Modal Assurance Criterion (MAC) values for all the obtained mode shapes are all over 0.95. Mode shapes 1, 3 and 6 have MAC values of more than 0.99, which indicates great similarities in mode shapes between the results from vision-based method and ground truth data. The accuracy and effectiveness of the proposed method for vibration displacement and modal identification are validated.

4.5 Structural damage identification

In this section, the above obtained natural frequencies are used for structural damage identification with the presented optimisation algorithm L-SHADE in Section 3. The implementation of this algorithm and the accuracy of damage identification are presented herein.

4.5.1 Initial model updating

For a better accuracy, spatial beam elements with two nodes are used to represent the truss element and build the finite element model, where each node has 6 degrees of freedom including three axial translational and three rotational degrees of freedom. Before conducting structural damage identification by using optimization algorithms, there exists some modelling errors between the finite element (FE) model and synthetic model of the truss structure. These deviations are manifested as discrepancies in natural frequencies, and should be minimised before the FE model is adopted for damage identification (Hao and Xia 2002). Therefore, the Young's modulus of each element of the analytical FE model inside the field of view of the camera are updated, such that the discrepancy between the frequencies is minimal.

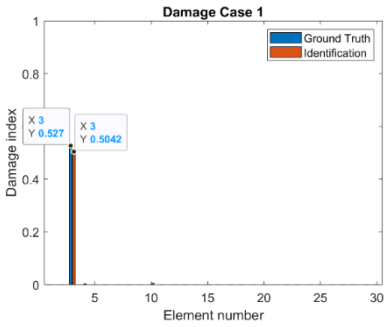
By minimising the discrepancies of the identified natural frequencies between the 'Undamaged Case' and the analytically calculated ones from the initial finite element model, an updated finite element model is constructed to best describe the vibration of the synthetic experimental model. Based on the same objective function as shown in Eq. (9), the six natural frequencies of the frame model under the undamaged status presented in Table 4 are set as the baseline for updating. All 29 elements shown in Fig. 4(b) are deemed as variables in the minimisation, and the initially assumed Young's modulus of 18.5 GPa is used as input of the initial value for every element. The L-SHADE algorithm introduced in Section 3 is used for initial model updating.

In the L-SHADE algorithm, parameters M_{CR} and M_F are set as 0.5 in the initial stage. Control rate p, memory size H and the maximum iteration are set as 0.11, 5 and 200, respectively. The population size shrinks from 18 to 4 linearly. The identified natural frequencies from the obtained displacement responses by using the vision-based method, and the updated finite element model are used together for damage identification with the presented L-SHADE algorithm in Section 3.

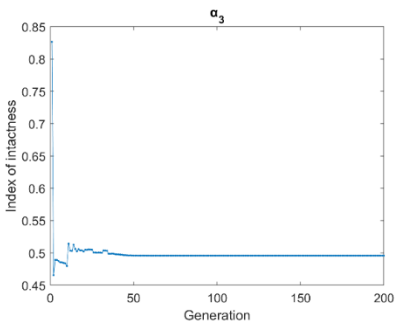
After model updating, the natural frequencies of the updated undamaged model are shown in Table 4. It is observed that the relative errors between natural frequencies obtained from measured displacement responses and updated FE model are very small. This updated finite element model, and its natural frequencies are used as the baseline for the subsequent structural damage identification in damage cases.

4.5.2 Damage localisation and severity assessment for Cases 1-5

For damage cases 1-5, the ground truth values of the damage information are provided. Therefore, the damage identification results of these five scenarios are presented first. The initial elemental stiffness indices are set to 1, before the 1st iteration when using the algorithm. The

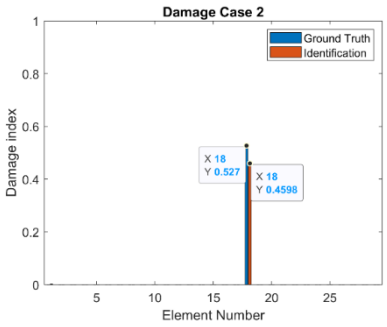


(a) Comparison of damage indices

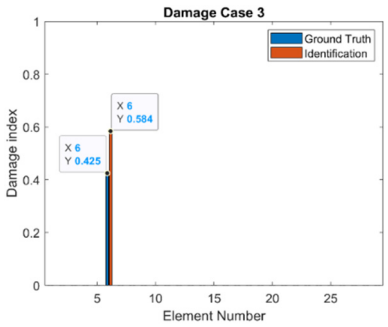


(b) Evolution process of stiffness parameter α3

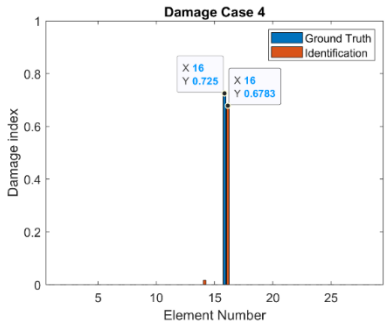
Fig. 11 Damage identification results for Case 1



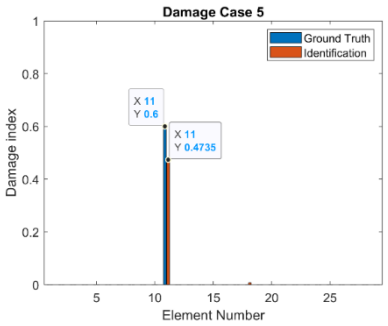
(a) Case 2



(b) Case 3



(c) Case 4



(d) Case 5

Fig. 12 Damage identification results for Cases 2 to 5

parameters of L-SHADE are set the same as those used in the initial model updating in Section 4.5.1. For these damage scenarios, there is only one component of the truss that is damaged. Therefore, if there is more than one damage index that is less than 1 (not intact), the element with the largest damage is identified as the damaged element, and the obtained damage index value is identified as its stiffness damage parameter.

The identification results for Case 1 are shown in Fig. 11. It can be observed that only the No.3 element shows a damage index value significantly larger than 0, and only the No.4 and No.10 elements which are near the damaged element show minor identified damage values, which are negligible. As shown in Fig. 11(a), the identified damage index for element No.3 is 0.5042, compared with the ground truth value 0.527. This shows that the damage is correctly located, and the severity is well identified. Fig. 11(b) shows the evolution process of elemental stiffness index α_3 takes around 40 iterations for the algorithm to reach a stable convergence near 0.4958, which indicates a damage index of 0.5042. In fact, with the introduction of sparsity regularisation item, the identified damage indices can converge faster and better. Several very minor false identification results are observed in the adjacent elements, such as elements No.4 and No.10 close to No. 3 in Case 1. However, the damage is identified accurately, demonstrating the accuracy of using the presented optimization algorithm for damage identification.

The identification results for the other cases, such as Cases 2 to 5, are shown in Fig. 12. All the results show that damages are located correctly. Only one obvious major damage is identified in every case, which suggests the existence of only one damage. Moreover, there is almost no false identification, and the results are close to the real values. Hence, the effectiveness and correctness of the proposed method for structural damage identification are demonstrated.

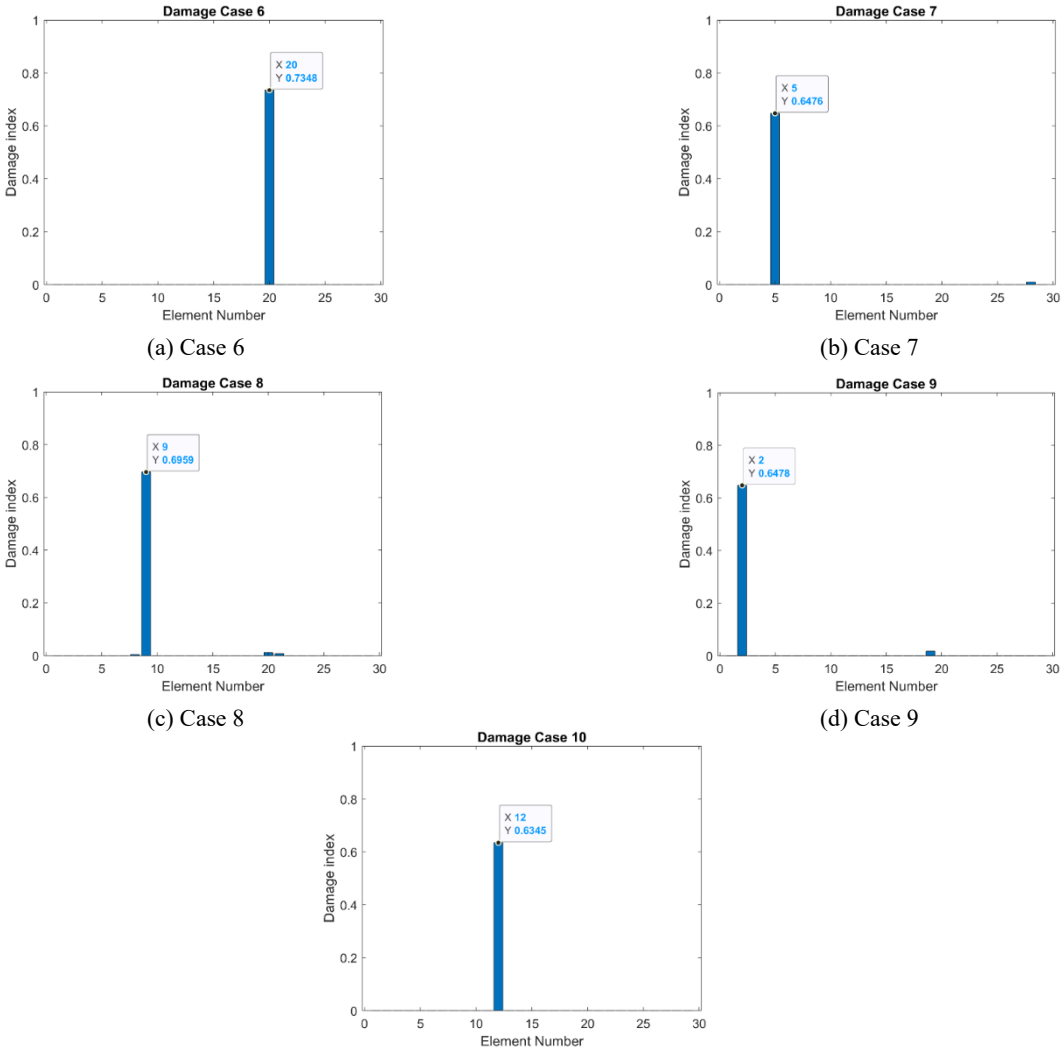


Fig. 13 Damage identification results for Cases 6 to 10

4.5.3 Damage localisation and severity identification for Cases 6-10

Damage results for Cases 6 to 10 are presented in this section. The displacement data are also extracted from the corresponding video clips, and natural frequencies identified from obtained displacement responses are used for modal identification and subsequent damage identification. The parameters for the L-SHADE algorithm are the same as those in Cases 1 to 5. The identification results are shown in Fig. 13. Since ground truth damage information is not provided, only the identified results in these cases are presented. Generally speaking, damage locations are clearly identified, only one critical damage occurs in each case. All the corresponding damage severities are greater than 0.4, which satisfies the damage conditions of every damage scenario. Overall, the damage of every damage case can be successfully identified, which shows the performance of the proposed approach in damage identification.

5. Conclusions

This study presents a novel motion tracking method for vibration displacement measurement and structural damage identification. The homography transformation is applied on the video from non-parallel cases, so that the scale factor on the front plane is uniform and the videos can be tracked directly. FAST feature detection with adaptive threshold and KLT are used to obtain stable corner feature points and track the motion of feature points. The accuracy of obtained vibration displacement responses is compared and verified with ground truth data. Natural frequencies and mode shapes are identified by SSI. High MAC values between the identified mode shapes from data of vision-based method and ground truth are obtained. Elemental damage is treated as reduction of Young's modulus, and an objective function is constructed. The Bayesian inference and $L_{0.5}$-norm sparse regularization are used to the initial numerical model update based on natural frequencies and subsequent damage identification for the truss structure under damage conditions. The Successive-history memory parameter adaption and Linear Population Reproduction strategy are

combined into DE. The newly built L-SHADE is used to perform the optimization of the objective function with a high dimension of parameters, for all the damage cases. The identification results show that the proposed method can effectively locate and identify the damage severity.

Acknowledgments

The first author disclosed receipt of the financial support by China Scholarship Council Grant No.201806380151. The corresponding author acknowledges the support from Australian Research Council Discovery project DP210103631, "AI Assisted Probabilistic Structural Health Monitoring with Uncertain Data".

References

Baqersad, J., Niezrecki, C. and Avitabile, P. (2015), "Extracting full-field dynamic strain on a wind turbine rotor subjected to arbitrary excitations using 3D point tracking and a modal expansion technique", *J. Sound Vib.*, **352**, 16-29. https://doi.org/10.1016/j.jsv.2015.04.026

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008), "Speeded-Up Robust Features (SURF)", *Comput. Vision Image Und.*, **110**(3), 346-359. https://doi.org/10.1016/j.cviu.2007.09.014

Brest, J. and Maučec, M. (2008), "Population size reduction for the differential evolution algorithm", *Appl. Intell.*, **29**, 228-247. https://doi.org/10.1007/s10489-007-0091-x

Brest, J., Greiner, S., Boskovic, B., Mernik, M. and Zumer, V. (2006), "Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems", *IEEE Transact. Evolut. Comput.*, **10**(6), 646-657. https://doi.org/10.1109/TEVC.2006.872133

Cawley, P. and Adams, R.D. (1979), "The location of defects in structures from measurements of natural frequencies", *J. Strain Anal. Eng.*, **14**(2), 49-57. https://doi.org/10.1243/03093247v142049

Chen, Z.P. and Yu, L. (2017), "A novel PSO-based algorithm for structural damage detection using Bayesian multi-sample objective function", *Struct. Eng. Mech., Int. J.*, **63**(6), 825-835. https://doi.org/10.12989/sem.2017.63.6.825

Chen, F., Chen, X., Xie, X., Feng, X. and Yang, L. (2013), "Full-field 3D measurement using multi-camera digital image correlation system", *Opt. Laser Eng.*, **51**(9), 1044-1052. https://doi.org/10.1016/j.optlaseng.2013.03.001

Das, S. and Suganthan, P.N. (2011), "Differential evolution: A survey of the state-of-the-art", *IEEE Transact. Evolut. Comput.*, **15**(1), 4-31. https://doi.org/10.1109/TEVC.2010.2059031

Ding, Z., Li, J. and Hao, H. (2019), "Structural damage identification using improved Jaya algorithm based on sparse regularization and Bayesian inference", *Mech. Syst. Signal Pr.*, **132**, 211-231. https://doi.org/10.1016/j.ymssp.2019.06.029

Feng, D. and Feng, M.Q. (2018), "Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review", *Eng. Struct.*, **156**, 105-117. https://doi.org/10.1016/j.engstruct.2017.11.018

Hao, H. and Xia, Y. (2002), "Vibration-based damage detection of structures by genetic algorithm", *J. Comput. Civil Eng.*, **16**(3), 222-229. https://doi.org/10.1061/(ASCE)0887-3801(2002)16:3(222)

Hao, H., Bi, K., Chen, W., Pham, T.M. and Li, J. (2023), "Towards next generation design of sustainable, durable, multi-hazard resistant, resilient, and smart civil engineering structures", *Eng. Struct.*, **277**, 115477. ttps://doi.org/10.1016/j.engstruct.2022.115477

Harris, C. and Stephens, M. (1988), "A combined corner and edge detector", *Proceedings of the 4th Alvey Vision Conference*, Manchester, UK, September.

Helfrick, M.N., Niezrecki, C., Avitabile, P. and Schmidt, T. (2011), "3D digital image correlation methods for full-field vibration measurement", *Mech. Syst. Signal Pr.*, **25**(3), 917-927. https://doi.org/10.1016/j.ymssp.2010.08.013

Hou, R., Xia, Y., Bao, Y. and Zhou, X. (2018), "Selection of regularization parameter for l1-regularized damage detection", *J. Sound Vib.*, **423**, 141-160. https://doi.org/10.1016/j.jsv.2018.02.064

Kuddus, M.A., Li, J., Hao, H., Li, C. and Bi, K. (2019), "Target-free vision-based technique for vibration measurements of structures subjected to out-of-plane movements", *Eng. Struct.*, **190**, 210-222. https://doi.org/10.1016/j.engstruct.2019.04.019

Lee, J. (2009), "Identification of multiple cracks in a beam using natural frequencies", *J. Sound Vib.*, **320**(3), 482-490. https://doi.org/10.1016/j.jsv.2008.10.033

Loop, C. and Zhang, Z. (1999), "Computing Rectifying Homographies for Stereo Vision", *Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1. https://doi.org/10.1109/CVPR.1999.786928

Lowe, D.G. (1999), "Object recognition from local scale-invariant features", *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Corfu, Greece, September.

Lu, Z.R. and Law, S.S. (2007), "Features of dynamic response sensitivity and its application in damage detection", *J. Sound Vib.*, **303**(1-2), 305-329. https://doi.org/10.1016/j.jsv.2007.01.021

Lu, Z.-R. and Wang, L. (2017), "An enhanced response sensitivity approach for structural damage identification: convergence and performance", *Int. J. Numer. Meth. Eng.*, **111**(13), 1231-1251. https://doi.org/10.1002/nme.5502

Lucas, B. and Kanade, T. (1981), "An Iterative Image Registration Technique with an Application to Stereo Vision", Pr*oceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver, B.C., Canada, August.

Narazaki, Y., Gomez, F., Hoskere, V., Smith, M.D. and Spencer Jr, B.F. (2020), "Efficient development of vision-based dense three-dimensional displacement measurement algorithms using physics-based graphics models", *Struct. Health Monitor.*, **20**(4), 1841-1863. https://doi.org/10.1177/1475921720939522

Pan, B., Qian, K., Xie, H. and Asundi, A. (2009), "Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review", *Measure. Sci. Technol.*, **20**(6), 062001. https://doi.org/10.1088/0957-0233/20/6/062001

Peeters, B. and De Roeck, G. (1999), "Reference-based stochastic subspace identification for output-only modal analysis", *Mech. Syst. Signal Pr.*, **13**(6), 855-878. https://doi.org/10.1006/mssp.1999.1249

Peng, F., Tang, K., Chen, G. and Yao, X. (2009), "Multi-start JADE with knowledge transfer for numerical optimization", *Proceedings of 2009 IEEE Congress on Evolutionary Computation*, Trondheim, Norway, May.

Qin, A.K., Huang, V.L. and Suganthan, P.N. (2009), "Differential evolution algorithm with strategy adaptation for global numerical optimization", *IEEE Transact. Evolut. Comput.*, **13**(2), 398-417. https://doi.org/10.1109/TEVC.2008.927706

Reu, P.L., Rohe, D.P. and Jacobs, L.D. (2017), "Comparison of DIC and LDV for practical vibration and modal measurements", *Mech. Syst. Signal Pr.*, **86**, 2-16. https://doi.org/10.1016/j.ymssp.2016.02.006

Rosten, E. and Drummond, T. (2005), "Fusing points and lines for high performance tracking", *Proceedings of IEEE International*

Conference on Computer Vision, Bejing, China, October.
Shao, Y., Li, L., Li, J., An, S. and Hao, H. (2021), "Computer vision based target-free 3D vibration displacement measurement of structures", *Eng. Struct.*, **246**, 113040. https://doi.org/10.1016/j.engstruct.2021.113040
Shi, J. and Tomasi, C. (1994), "Good features to track", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, June.
Storn, R. (1996), "On the usage of differential evolution for function optimization", *Proceedings of North American Fuzzy Information Processing*, Berkeley, California, USA, June.
Storn, R. and Price, K. (1997), "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces", *J. Global Optim.*, **11**(4), 341-359. https://doi.org/10.1023/A:1008202821328
Tan, D., Li, J., Hao, H. and Nie, Z. (2023), "Target-free vision-based approach for modal identification of a simply-supported bridge", *Eng. Struct.*, **279**, 115586. https://doi.org/10.1016/j.engstruct.2022.115586
Tanabe, R. and Fukunaga, A. (2013), "Success-history based parameter adaptation for Differential Evolution", *Proceedings of 2013 IEEE Congress on Evolutionary Computation*, Cacun, Mexico, June.
Tanabe, R. and Fukunaga, A.S. (2014), "Improving the search performance of SHADE using linear population size reduction", *Proceedings of 2014 IEEE Congress on Evolutionary Computation*, Beijing, China, July.
Tikhonov, A.N. (1963), "On the solution of ill-posed problems and the method of regularization", *Dokl. Akad. Nauk SSSR*, **151**(3), 501-504.
Titurus, B. and Friswell, M.I. (2008), "Regularization in model updating", *Int. J. Numer. Meth. Eng.*, **75**(4), 440-478. https://doi.org/10.1002/nme.2257
Tomasi, C. and Kanade, T. (1999), "Shape and Motion from Image Streams: a Factorization Method—Part 3 Detection and Tracking of Point Features", *Int. J. Comput. Vision*, **9**(2), 137-54. https://doi.org/10.1007/BF00129684
Trajković, M. and Hedley, M. (1998), "Fast corner detection", *Image Vision Comput.*, **16**(2), 75-87. https://doi.org/10.1016/S0262-8856(97)00056-5
Vlašić, I., Đurasević, M. and Jakobović, D. (2019), "Improving genetic algorithm performance by population initialisation with dispatching rules", *Comput. Ind. Eng.*, **137**, 106030. https://doi.org/10.1016/j.cie.2019.106030
Xu, Z., Zhang, H., Wang, Y., Chang, X. and Liang, Y. (2010), "L1/2 regularization", *Science China Information Sciences*, **53**(6), 1159-1169. https://doi.org/10.1007/s11432-010-0090-0
Yoon, H., Elanwar, H., Choi, H., Golparvar-Fard, M. and Spencer Jr, B.F. (2016), "Target-free approach for vision-based structural system identification using consumer-grade cameras", *Struct. Control Health Monitor.*, **23**(12), 1405-1416. https://doi.org/10.1002/stc.1850
Zepeng, C. and Yu, L. (2017), "A novel PSO-based algorithm for structural damage detection using Bayesian multi-sample objective function", *Struct. Eng. Mech., Int. J.*, **63**, 825-835. https://doi.org/10.12989/sem.2017.63.6.825
Zhang, Z. (2003), "Single-view geometry of a rectangle with application to whiteboard image rectification", Microsoft Research; Redmond, WA, USA.
Zhang, J. and Sanderson, A.C. (2009), "JADE: adaptive differential evolution with optional external archive", *IEEE Transact. Evolut. Comput.*, **13**(5), 945-958. https://doi.org/10.1109/TEVC.2009.2014613
Zhang, C.D. and Xu, Y.L. (2016), "Comparative studies on damage identification with Tikhonov regularization and sparse regularization", *Struct. Control Health Monitor.*, **23**(3), 560-579. https://doi.org/10.1002/stc.1785
Zhong, S. and Oyadiji, S.O. (2011), "Detection of cracks in simply-supported beams by continuous wavelet transform of reconstructed modal data", *Comput. Struct.*, **89**(1-2), 127-148. https://doi.org/10.1016/j.compstruc.2010.08.008