Simultaneous measurement of multipoint mechanical vibration for thin-walled parts using monocular machine vision optical flow tracking

Jigang Wu*, Jun Shao
Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, P.R. China

Abstract: Thin-walled parts (TWPs) have been widely employed in various industry fields and put forward a specific requirement of simultaneous multipoint vibration measurement for their dynamic behaviours understanding. The development of machine vision theory and technology enable the machine vision-based technique to measure the vibration of mechanical equipment. Lucas-Kanade based template tracking (LKBTT) can provide a robust, longer term motion estimation of the interesting object, but lacks spatial resolution and also deeper use of contextual cues. Lucas-Kanade based optical flow tracking (LKBOPT) is usually used for frame-by-frame feature-level motion estimation, and has the potential to satisfy the vibration measurement requirements of TWPs. LKBOPT has not been introduced to vibration measurement yet, and also had the challenge of long-term motion estimation. In order to satisfy the simultaneous multipoint vibration measurement requirements of TWPs, a coarse-to-fine tracking method is developed by introducing LKBOPT to vibration measurement and combining it with LKBTT. A region of interest detection algorithm is investigated to acquire the tracking template according to the measurable object. LKBTT is introduced to realize the coarse tracking and obtain the vibration information of the tracking template. An improved LKBOPT is investigated to realize the fine tracking and acquire the vibration information of multipoint inside the tracking template simultaneously. Finally, the simultaneous measurement system of multipoint mechanical vibration for TWPs is utilized with a high-speed camera system based on both Python and open source computer vision (OpenCV) libraries. The vibration measurement performance of the proposed method has been experimentally verified with two experiments under laboratory conditions by comparing the results with those using accelerometers. The experiment has shown that the proposed method is effective in the remote simultaneous multipoint vibration measurement of TWPs with satisfactory accuracy and efficiency.

Keywords: Simultaneous measurement; multipoint mechanical vibration; thin-walled parts; monocular machine vision; optical flow tracking

1. Introduction

Thin-walled parts (TWPs) have been widely employed in various industry fields, especially in aerospace, automobile and energy industry owing to their lightweight and high specific strength properties [1-2]. However, TWPs are apt to deform and vibrate excessively during the running because of the low rigidity. The over deformation and vibration will result in fatigue damage and even breakdown after long-time running. Therefore, it is very necessary to understand the dynamic behaviours of TWPs for dynamic optimization, structural damage identification, etc. Furthermore, TWPs are flexible bodies and the vibration characteristic is different from that of rigid bodies which vibrate all in the same way. So, the simultaneous multipoint vibration measurement is the specific requirement of TWPs as to the rigid body.

Mechanical vibration measurement is essential to understand the dynamic behaviours of a mechanical equipment [3], it also serves as the basis for various engineering practices such as natural frequency or
resonant frequency estimation [4]. Traditional contact-type measurement devices, such as accelerometers, are widely used to collect vibration information [5-6]. However, this contact-type measurement approach has the deficiencies of mass loading effect, it is particularly problematic for thin-walled parts, low spatial sensing resolution, costly, time and labour-consuming installation, and so on. Therefore, significant interest has been shown in the development of non-contacting vibration measurement techniques [7].

With the development of machine vision theory and technology, machine vision-based measurement approaches which extract motion information from video image sequences, have become realistic alternatives among all non-contacting vibration measurement approaches recently, bringing with the advantages of remote measurement, high spatial resolution, relatively low cost and flexible in application [8-10]. Machine vision-based measurement approaches have been widely applied in both research and industrial applications. Son et al. [11] proposed a fast high-resolution vibration measurement method based on vision technology for structure. And they developed the first- and second-order derivative algorithms to effectively detect the features of the measurement points in the image. Figueiredo et al. [12] presented a new methodology to measure natural frequencies of aeronautical structures using a computer vision system. Qiu et al. [13] developed a novel non-contact vibration measurement method using binocular vision sensors for piezoelectric flexible hinged plate, their experimental results demonstrate that the binocular vision measurement method is feasible on the low frequency vibration of the piezoelectric flexible hinged plate. Aforementioned approaches always use a series of image processing techniques to perform video vibration information extraction, hereby called them as image processing-based measurement approaches.

Although, image processing-based measurement approaches have gained relatively satisfactory results, they still have some intrinsic disadvantages. First of all, these approaches typically require user markings and structure edges as the features on the surface. Several applications without markers have been developed recently. Wu et al. [14] developed a non-contact, unmarked computer vision measurement method, they also analyzed the validity of the proposed method for the two-dimensional (2D) vibration displacement of hoisting vertical ropes. Kuddus et al. [15] proposed a target-free vision-based technique for measuring vibration displacement, and it is not required artificial targets on structures for displacement measurement. However, the measurable objects are limited by the need for the image to contain specific features in these approaches. Furthermore, these approaches are also based on edge detection, image segmentation, image correlation, etc., to obtain actual motion information in which many parameters need to be adjusted simultaneously to adapt to the specific type of image conditioning, the robustness of these approaches reduces remarkably. Finally, these approaches are complex in terms of processing requirements and the quantity of data acquiring processing is huge, the computational efficiency cannot satisfy the requirements of field measurement.

Optical flow estimation, a fundamental task in machine vision, is the process of finding a displacement vector for each pixel (called dense optical flow) or each feature (called sparse optical flow) between two consecutive images in a video image sequence, and can date back to the work of Horn and Schunck [16], and of Lucas and Kanade [17]. It has attracted significant research effort, and made considerable progress in recently years [18]. In generally, the applications only require focus on an object of interest in the image, and only a subset region of motion vectors is required. Therefore, the Lucas-Kanade algorithm has been investigated and extended the more widely. A unifying framework for the Lucas-Kanade algorithm is summarised [19].

Lucas-Kanade based template tracking (LKBTT) finds the closest matching region in consecutive frames. LKBTT was introduced to vibration measurement to eliminate the problems of image processing-based approaches [20-26]. LKBTT can provide a robust, longer term motion estimation featuring a holistic description of the interesting object. However, it still has some intrinsic shortcomings. Firstly, only a specific and unique region of content of the image (e.g. a screw head) can be used as the template, therefore
it can only provide sparse and discrete measurements. Secondly, the size of the template is a little big which will inevitably result in low resolution spatial sensing. Finally, only the motion information of the entire template can be obtained by this approach, the motion information of the interesting object or specific multipoint inside the template cannot be extracted. In conclusion, LKBTT cannot satisfy the vibration measurement requirements of TWPs.

Lucas-Kanade based optical flow tracking (LKBBOPT) is usually applied for frame-by-frame feature-level motion estimation, which means that LKBBOPT can take feature points as tracking object. It has the potential to satisfy the vibration measurement requirements of TWPs, nevertheless it has not been introduced to vibration measurement yet. Furthermore, LKBBOPT has been widely used as a motion feature tracking, long-term motion estimation is still a challenge for optical flow estimation even when modern optical flow approaches are employed \[27\], because it is based on brightness constancy assumption.

In this case, introducing LKBBOPT to vibration measurement and combining it with LKBTT would be useful to satisfy the simultaneous multipoint vibration measurement requirements of TWPs. As a result, a coarse-to-fine tracking method is developed, and a simultaneous measurement of multipoint mechanical vibration for TWPs using monocular machine vision optical flow tracking is proposed. An interesting object detection algorithm is investigated to acquire the image region of the measured object from an initial frame of the video image sequence and then the image region is taken as the tracking template. LKBTT is introduced to achieve the coarse tracking and obtain the motion information of the tracking template. An improved LKBBOPT is investigated to realize the fine tracking and provide the motion information of the multipoint inside the tracking template region. Finally, the simultaneous measurement system of multipoint mechanical vibration for TWPs is utilized with a high-speed camera system based on both Python and open source computer vision (OpenCV) libraries. The vibration measurement performance of the proposed method has been experimentally verified with two experiments under laboratory conditions by comparing the results with those using accelerometers. The experiment has shown that the proposed method is effective in the remote simultaneous multipoint vibration measurement of TWPs with satisfactory accuracy and efficiency.

This paper is organized as follows: the theory and method of the proposed simultaneous multipoint vibration measurement system are presented in Section 2. Two experiments for performance verification are presented in Section 3, and conclusions are drawn in Section 4.

2. Theory and method

The theory and method of the proposed simultaneous multipoint vibration measurement system are introduced in this section.

The flowchart of the proposed simultaneous multipoint vibration measurement method is shown in Fig. 1. Compared with LKBTT vibration measurement approach, the proposed simultaneous multipoint vibration measurement method has some advantages. Firstly, the proposed method can realize simultaneous multipoint vibration measurements for TWPs. Secondly, the size of the template is decided by the detected object of interest which provides contextual cues for LKBBOPT. Finally, the motion information of the object (or specific multipoint inside the template) can be extracted.
2.1 Coarse-to-fine tracking principle

The basic procedure of the coarse-to-fine tracking method is shown in Fig. 2, where Fig. 2 (a) depicts the initial frame of the video image sequence. The object of interest is detected and then taken as the tracking template. The features inside the template are detected in accordance with the requirement of improved LKBOPT. Fig. 2 (b) shows the n-th frame of the video image sequence. The large displacement of the whole template is measured with LKBTT. The small displacements of multipoint inside the template are measured with improved LKBOPT.

If we suppose that a motion vector $\mathbf{f}_p = (u_p, v_p)$ is attached to a pixel $\mathbf{p} = (x, y)$, given the initial frame $I(x, y, 0)$ and the n-th frame $I(x, y, t)$, then hierarchy of memberships of the $M$ layers to which a pixel belongs is also established. In this case, $\mathbf{p} \in S_i$ with $i = 1, \ldots, M$ and $S_1 \subset S_2 \subset S_3 \subset \ldots$, and subsequently $S_{M-1} \subset S_M$. All of the layers $S_i$ are related to meaningful regions which usually denote some level of semantic and can be considered as a sub-set of an image. Supposing each one of these layers $S_i$ has an assigned motion vector $\mathbf{w}_i$ at location $\mathbf{p}$ with respect to its parent in the hierarchy, the total motion of the pixel $\mathbf{p}$ can be expressed as follows:

$$\mathbf{f}_p = \mathbf{v}_p + \sum_{i=1}^{M} \mathbf{w}_i$$  \hspace{1cm} (1)
where \( \varepsilon_p \) is the motion vector of the region or the whole object containing the point \( p \).

In this case, the motion computation of each layer in the hierarchy is supported by the calculation of the predecessor layer. For most cases, a simple scheme with two layers will be sufficient to obtain accurate tracking result. Thus, simplifying the Eq. (1), we can write:

\[
f_p = \varepsilon_p + w_o
\]  

(2)

where \( w_o \) is the differential motion of the point \( p \) with respect to the motion of the entities the point belongs to, \( \varepsilon_p \) will be obtained with LKBTT and \( w_o \) will be acquired by improved LKBOPT.

### 2.2 Template detection algorithm

The task of the template detection algorithm is to detect the measured object within initial fame of the video image sequence and to draw the bounding rectangle with minimum area around it. The region contained inside the bounding rectangle of the image is then taken as the tracking template. The flowchart of the template detection algorithm is shown in Fig. 3. The total algorithm was coded using OpenCV functions and the Python. The details can refer to the OpenCV-Python Tutorials.

![Flowchart of the template detection algorithm](image)

**Fig. 3 Flowchart of the template detection algorithm**

### 2.3 LKBTT algorithm

If \( I_i(x) \) represents the \( i \)th image in a given video image sequence, where \( x = (x, y)^T \) is a column vector containing the pixel coordinates and \( i = 0, 1, 2, L \) represents the frame number, then the template \( T(x) \) is an extracted sub-region of the initial frame \( I_0 \), which contains the object of interest to be tracked in the video image sequence. LKBTT algorithm must match the template image \( T(x) \) with the subsequent frames \( I_i(x) \) and then extract the motion information of the object of interest.

The matching of images in the video image sequence is parameterized as a warp function \( W(x;p) \). This function transforms the pixel coordinate \( x \) of the template \( T \) to the sub-pixel coordinate \( W(x;p) \) of the subsequent image \( I_i; \ p = (p_1, L, p_n)^T \) represents a vector of transformation parameters of the warp function. The objective of template-based tracking is to search for the optimized parameters \( p \) so that the
warped image \( I_i(W(x;p)) \) matches the template \( T(x) \) as closely as possible \(^{[17]}\).

In LKBTT algorithm, the warp function plays a very important role in tracking different types of motion. For instance, 2-Dimension affine transformation is widely used in general tracking. It consists of six independent parameters and takes many linear transformations into account, including translation, rotation, shear mapping, and scaling. In order to improve the algorithm performance, the warp function \( W(x;p) \) is simplified to \( p = (p_x, p_y)^T \) in Reference \(^{[22]}\), and only comprises two independent parameters since the authors believe that the movements of structures are mostly in-plane motions with a limited range, and deformation and rotation can be neglected in a short capture. This simplification is sometimes appropriate for application to large-scale structures, but it has limitations for measurement of TWPs, so the un-simplified warp function \( W(x;p) \) has been used in the proposed simultaneous multipoint vibration measurement method.

### 2.4 Improved LKBOPT algorithm

If \( I \) and \( J \) are two sequential grayscale images, \( I(x) \) and \( J(x) \) are the grayscale values of the two sequential images at the location \( x = (x, y)^T \); Image \( I \) is referred to the first image, and image \( J \) as the second. Supposing \( n_x \) and \( n_y \) are the width and height of the two images, then the lower right pixel coordinate vector is \( (n_x - 1, n_y - 1)^T \).

Supposing \( u = (u_x, u_y)^T \) is an image point on the first image \( I \), the goal of optical flow estimation is to find the location \( v = u + d = (u_x + d_x, u_y + d_y)^T \) on the second image \( J \) so that \( I(u) \) and \( J(v) \) are as similar as possible. Supposing the vector \( d = (d_x, d_y)^T \) is the image velocity at \( u \), i.e. the optical flow at \( x \), then because of the problem of aperture, it is necessary to define the notion of similarity in a neighborhood sense. If \( \omega_x \) and \( \omega_y \) are two integers, the residual function \( \varepsilon(d) \) can be defined as follows:

\[
\varepsilon(d) = \varepsilon(d_x, d_y) = \sum_{x=-\omega_x}^{u_x+\omega_x} \sum_{y=-\omega_y}^{u_y+\omega_y} [I(x, y) - J(x + d_x, y + d_y)]^2
\]

where the similarity function is measured on an image neighborhood of size \( (2\omega_x + 1) \times (2\omega_y + 1) \). Typical values for \( \omega_x \) and \( \omega_y \) are \( 2, 3, 4, 5, 6, 7 \) pixels\(^{[28]}\). It is preferable to have \( d_x \leq \omega_x \) and \( d_y \leq \omega_y \) in Eq. (3). A tradeoff exists between local accuracy and robustness when choosing the integration window size.

The initial guess of the flow vector \( g^L \) is used to translate the image patch in the second image \( J \). The optical flow solution is then available after the finest resolution optical flow computation:

\[
d = g^0 + d^0
\]
This solution may be expressed in the following extended form:

\[ d = \sum_{L=0}^{L_m} 2^L d^L \]  

(5)

The clear advantage of a pyramidal implementation is that each residual optical flow vector \( d^L \) can be kept very small whilst computing a large overall pixel displacement vector \( d \). Assuming that each elementary optical flow computation step can handle pixel motions up to \( d_{\text{max}} \), then the overall pixel motion that the pyramidal implementation can handle becomes

\[ d_{\text{max final}} = (2^{L_m+1} - 1)d_{\text{max}}. \]

For example, for a pyramid depth of \( L_m = 3 \), means a maximum pixel displacement gain of 15. This enables large pixel motions to be resolved, while keeping the size of the integration window relatively small.

At every level \( L \) in the pyramid, the goal is finding the vector \( d^L \) that minimizes the matching function \( \varepsilon^L \) defined as follows.

\[
\varepsilon^L(d^L) = \varepsilon^L(d^L_x, d^L_y) = \sum_{x=a_x}^{a_x+\omega_x} \sum_{y=a_y}^{a_y+\omega_y} [I^L(x,y) - J^L(x + g^L_x, y + g^L_y + d^L_y)]^2
\]

(6)

Since the same type of operation is performed for all levels \( L \), the superscript \( L \) is dropped and the new images \( A \) and \( B \) are defined as follows [28]:

\[
\forall (x,y) \in [p_x - \omega_x - 1, p_x + \omega_x + 1] \times [p_y - \omega_y - 1, p_y + \omega_y + 1],
\]

\[ A(x,y) = B I^L(x,y) \]  

(7)

\[
\forall (x,y) \in [p_x - \omega_x, p_x + \omega_x] \times [p_y - \omega_y, p_y + \omega_y],
\]

\[ B(x,y) = B J^L(x + g_x^L, y + g_y^L) \]  

(8)

The domains of definition of \( A(x,y) \) and \( B(x,y) \) are slightly different. \( A(x,y) \) is defined over a window of size \( (2\omega_x + 3) \times (2\omega_y + 3) \) instead of \( (2\omega_x + 1) \times (2\omega_y + 1) \). This difference will become clear when computing spatial derivatives of \( A(x,y) \) using the central difference operator. For clarity purposes, the name of the displacement vector \( \vec{v} = (v_x, v_y)^T = d^L \) is also changed, as well as the image position vector \( p = (p_x, p_y)^T = u^L \). Following these changes, the goal is now to find the displacement vector \( \vec{v} = (v_x, v_y)^T \) that minimizes the matching function:

\[
\varepsilon(\vec{v}) = \varepsilon(v_x, v_y) = \sum_{x=p_x-\omega_x}^{p_x+\omega_x} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} [A(x,y) - B(x + v_x, y + v_y)]^2
\]

(9)

At the optimum, the first derivative of \( \varepsilon \) with respect to \( \vec{v} \) is zero:
\[
\frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \bigg|_{\mathbf{v} = \mathbf{v}_0} = [0 \ 0]
\] (10)

After expansion of the derivative:
\[
\frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} = -2 \sum_{x=p_x-\omega_y}^{p_x+\omega_y} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} (A(x, y) - B(x+v_x, y+v_y)) \left[ \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \right]
\] (11)

\( B(x+v_x, y+v_y) \) is substituted by its first order Taylor expansion about the point \( \mathbf{v} = [0 \ 0]^T \):
\[
\frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \approx -2 \sum_{x=p_x-\omega_y}^{p_x+\omega_y} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} (A(x, y) - B(x, y) - \left[ \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \right]^T [\mathbf{v}]) \left[ \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \right]
\] (12)

The quantity \( A(x, y) - B(x, y) \) can be interpreted as the temporal image derivative at the point \( [x \ y]^T \):
\[
\nabla(x, y) \in [p_x - \omega_x, p_x + \omega_x] \times [p_y - \omega_y, p_y + \omega_y],
\]
\[
\delta I(x, y) B A(x, y) - B(x, y)
\] (13)

The matrix \( \left[ \frac{\partial B}{\partial x} \frac{\partial B}{\partial y} \right] \) is merely the image gradient vector, and the notation can be changed as follows:
\[
\nabla I = \left[ \begin{array}{c} I_x \\ I_y \end{array} \right] B \left[ \begin{array}{cc} \frac{\partial B}{\partial x} & \frac{\partial B}{\partial y} \end{array} \right]^T
\] (14)

The image derivatives \( I_x \) and \( I_y \) can be computed directly from the first image \( A(x, y) \) in the \((2\omega_x + 1) \times (2\omega_y + 1)\) neighborhood of the point \( p \) independently from the second image \( B(x, y) \). If a central difference operator is used for the derivative, the two derivative images have the following expressions:
\[
\forall (x, y) \in [p_x - \omega_x, p_x + \omega_x] \times [p_y - \omega_y, p_y + \omega_y],
\]
\[
I_x(x, y) = \frac{\partial A(x, y)}{\partial x} = \frac{A(x+1, y) - A(x-1, y)}{2}
\] (15)
\[
I_y(x, y) = \frac{\partial A(x, y)}{\partial y} = \frac{A(x, y+1) - A(x, y-1)}{2}
\] (16)

In practice, the Sharr operator is used for computing image derivatives\[^{[28]}\].

Eq. (12) can then be written:
\[
\frac{1}{2} \frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \approx \sum_{x=p_x-\omega_y}^{p_x+\omega_y} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} (\nabla I^T \mathbf{v} - \delta I) \nabla I^T
\] (17)
\[
\frac{1}{2} \frac{\partial \mathbf{e}(\mathbf{v})}{\partial \mathbf{v}} \approx \sum_{x=p_x-\omega_y}^{p_x+\omega_y} \sum_{y=p_y-\omega_y}^{p_y+\omega_y} \left[ \begin{array}{cc} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{array} \right] ^T \mathbf{v} - \left[ \begin{array}{c} \delta I I_x \\ \delta I I_y \end{array} \right]
\] (18)

If we denote:
\[ G = \sum_{x=p_x-\omega x, y=p_y-\omega y}^{p_x+\omega x, p_y+\omega y} \left[ I_x^2 I_y^2 I_x I_y \right] \quad \text{and} \quad \delta \bar{b} = \sum_{x=p_x-\omega x, y=p_y-\omega y}^{p_x+\omega x, p_y+\omega y} \left[ \delta I_x \delta I_y \right] \] (19)

Then Eq. (18) can be rewritten:
\[
\frac{1}{2} \frac{\partial \epsilon(\delta v)}{\partial \delta v} \approx G\delta v - \delta \bar{b}
\] (20)

Therefore, according to Eq. (18), the optimum optical flow vector is as follows:
\[
\delta v_{opt} = G^{-1}\delta \bar{b}
\] (21)

This expression is valid only if the matrix \( G \) is invertible; that is equivalent to saying that the image \( A(x, y) \) contains gradient information in both \( x \) and \( y \) directions in the neighborhood of the point \( p \), and this is the requirement that the tracking features need to satisfy.

This is the standard Lucas-Kanade optical flow equation, which is valid only if the pixel displacement is small. In practice, to get an accurate solution, it is necessary to iterate multiple times on this scheme.

Supposing \( k \) is the iterative index, initialized to 1 at the very first iteration. The algorithm is described recursively as follows: at a generic iteration \( k \geq 1 \), assume that the previous computations from iterations \( 1, 2, L, k-1 \) provide an initial guess \( \delta v^{k-1} = (v_x^{k-1}, v_y^{k-1})^T \) for the pixel displacement \( \delta v \). Supposing \( B_k \) is the new translated image according to that initial guess \( \delta v^{k-1} \):
\[
\forall (x, y) \in [p_x-\omega x, p_x+\omega x] \times [p_y-\omega y, p_y+\omega y], \quad B_k(x, y) = B(x + v_x^{k-1}, y + v_y^{k-1})
\] (22)

The goal is then to compute the residual pixel motion vector \( \delta \eta = [\eta_x, \eta_y]^T \) that minimizes the error function
\[
\epsilon^k(\delta \eta) = \epsilon(\delta \eta_x, \delta \eta_y) = \sum_{x=p_x-\omega x, y=p_y-\omega y}^{p_x+\omega x, p_y+\omega y} [A(x, y) - B_k(x + \eta_x, y + \eta_y)]^2
\] (23)

The solution of this minimization can be computed through a one-step Lucas-Kanade optical flow computation (Eq. (11))
\[
\delta \eta^k = G^{-1}\delta \bar{b}_k
\] (24)

where the \( 2 \times 1 \) vector \( \delta \bar{b}_k \) is defined as follows:
\[
\delta \bar{b}_k = \sum_{x=p_x-\omega x, y=p_y-\omega y}^{p_x+\omega x, p_y+\omega y} \left[ \delta I_x(x, y) I_x(x, y) \right]
\] (25)

where the \( k \)th image difference \( \delta I_k \) is defined as follows:
\[
\forall (x, y) \in [p_x-\omega x, p_x+\omega x] \times [p_y-\omega y, p_y+\omega y],
\quad \delta I_k = A(x, y) - B_k(x, y)
\] (26)
The spatial derivatives $I_x$ and $I_y$ (at all points in the neighborhood of $\vec{p}$) are computed only once at the beginning of the iterations following Eq. (15) and (16). Therefore the $2 \times 2$ matrix $G$ remains constant throughout the iteration loop. This constitutes a clear computational advantage.

The only quantity that needs to be recomputed at each step $k$ is the vector $\vec{h}_k$ that captures the amount of residual difference between the image patches after translation by the vector $\vec{v}^{k-1}$. Once the residual optical flow $\vec{\eta}_k$ is computed via Eq. (24), a new pixel displacement guess $\vec{v}_k$ is computed for the next iterative step $k+1$:

$$\vec{v}_k = \vec{v}^{k-1} - \vec{\eta}^k$$  \hspace{1cm} (27)

The iterative scheme goes on until the computed pixel residual $\vec{\eta}^k$ is smaller than a pre-set threshold or a maximum number of iterations is reached. On average, 5 iterations are enough to reach convergence \cite{28}. At the first iteration ($k=1$) the initial guess is initialized to zero:

$$\vec{v}^0 = [0 \ 0]^T$$  \hspace{1cm} (28)

Assuming that $K$ iterations were necessary to reach convergence, the final solution for the optical flow vector $\vec{v} = \mathbf{d}^L$ is:

$$\vec{v} = \mathbf{d}^L = \vec{v}^K = \sum_{k=1}^{K} \vec{\eta}^k$$  \hspace{1cm} (29)

This vector minimizes the error function described in Eq. (9). This ends the description of the iterative Lucas-Kanade optical flow computation. The vector $\mathbf{d}^L$ is fed into Eq. (5) and this overall procedure is repeated at all subsequent levels $L-1, L-2, L, 0$.

The improvements made to LKBOPT can be summarised as follows. Firstly, the complicated pyramidal implementation is not necessary in the proposed approach. The complicated pyramidal implementation of LKBOPT is needed to solve the large displacement problem, but because the large displacement is measured with LKBTT algorithm, the displacement between template $T^0$ in initial image $I^0$ and tracked template $T^i$ in current image $I^i$ is very small. Secondly, the optical flow estimation is calculated between template $T^0$ and template $T^i$, rather than between the previous image $I$ and current image $J$. The improvements enhance the robustness of the proposed method and solves the problem of long-term motion estimation.

3. Experiment and results

In this section, the experiment rig is built and two experiments are performed under laboratory conditions to verify the performance of the proposed simultaneous multipoint vibration measurement system.

3.1 The proposed simultaneous multipoint vibration measurement system
As shown in Fig. 4(a), the developed simultaneous multipoint vibration measurement system consists of a notebook computer connected to a video camera with a telescopic lens. As shown in Fig. 4(b), the telescopic lens has a large optical zoom capability and can be adjusted appropriately to capture mechanical equipment vibration at different distances. The video camera uses a CMOS sensor as the image receiver and can capture 8-bit grayscale images at high speed, which are streamed into the notebook computer through a USB 3.0 interface. The camera frame rate can reach up to 600 fps when the image resolution set to 640×480 pixels. The accuracy and capacity of the proposed system is also affected by the performance of computer, so the specifications of the laptop computer are provided. Technical specifications of the system are shown in Table 1.

The proposed algorithms have been integrated into the developed simultaneous multipoint vibration measurement system based on both Python and OpenCV libraries. The programmed software consists of calibration module and video analysis module. The calibration module is used to calculate the actual size that single pixel occupies on the target, and the parameters pixel size and focal length will be used. Any video image sequence captured by the high-speed video camera can be transferred to the video analysis module, and the motion information is then extracted by the proposed algorithms.

![Fig. 4 Simultaneous multipoint vibration measurement system: (a) experimental setup, (b) high-speed video camera](image)

| Tab. 1 Technical specifications of the developed simultaneous multipoint vibration measurement system |
|---|---|---|
| Component | Model | Technical Specification |
| Video camera | IMI Tech/IMB-3213UP | Image sensor type: 1/3 CMOS  
Maximum resolution: 640×480 pixel  
Real frame rate: 600 fps  
Image: 8-bit grayscale  
Pixel size: 4.8×4.8 μm  
Lens mount: C/CS-Mount  
Digital interface/Transfer rate: USB3.0/5Gbps  
S/N ratio: 40dB or better  
Dynamic range: 60dB in global shutter mode |
| Optical lens | TOKINA/TC1214-3MPG | Focal length: 12mm  
Aperture range: F1.4-16  
Mount: C-mount  
Angle of view: 39.6°×30.2° |
| Laptop computer | DELL/Inspiron 7447 14P-3948 | Intel(R) Core(TM) i7-4710HQ CPU @2.50 GHz  
8G RAM  
NVIDIA GeForce GTX 850M  
256G SSD + 1T HDD |
| Accessories | | Tripod, USB3.0 cable, etc. |
3.2 Vibration measurement accuracy and capacity verification experiment

The vibration measurement accuracy and capacity of the proposed system was first evaluated through a shaker vibration measurement experiment. In this experiment the proposed system was experimentally evaluated against a traditional accelerometer sensor. The vibration was recorded simultaneously by the accelerometer and the high-speed camera system to permit direct comparison. The experimental rig is shown in Fig. 5(a). An accelerometer (Model CA-YD-185 by SINOCERA Piezotronics Inc.) was fixed to a modal shaker (Model JZK-5T by SINOCERA Piezotronics Inc.). The shaker was driven by a signal generator (Model YE1311ET by SINOCERA Piezotronics Inc.) and the frequency and amplitude of the driving signal were set manually. The accelerometer vibration signal was sampled with a dynamic measuring (Model YE6231 by SINOCERA Piezotronics Inc.) and streamed into a computer through USB cable where it was recorded by general control & analysis software (Model YE7600 by SINOCERA Piezotronics Inc.). In parallel, the vibration of the accelerometer sensor was also recorded with the high-speed camera. The video images captured by the camera were digitized into 640×480 pixel images in 8-bit grey scales and also streamed into the computer through USB interface. The entire experimental setup is shown in Fig. 5(b). The sampling frequency of the general control & analysis software was 375Hz, and the frame rate of the camera was set to 375 fps. The sampling time for both devices was 16 seconds. The motion information of the accelerometer sensor was extracted using the proposed system on the attached computer. An image captured by the camera is displayed in Fig. 5(c), in which the red box is selected as the template and the green points are selected as the tracking features.

To evaluate the measurement capability from low frequency to high frequency, the vibration frequency of the shaker was set at successive values of 1Hz, 4Hz, 7Hz, 10Hz, 30Hz, 50Hz, 70Hz, 90Hz, 110Hz, 130Hz and 150Hz all at an amplitude of approximately 6mm. Since the frequency is set by an adjusting knob, the frequency cannot be fixed precisely at a certain value, the exact frequency being delivered was obtained by processing the vibration signal from the accelerometer. The displacement signal was acquired by the proposed system, whilst acceleration signal was obtained by the accelerometer. The acceleration signal was converted into its corresponding displacement for comparison. After conversion, both of the displacement signals were processed by FFT, and the amplitude frequency spectrums were plotted, as shown in Fig. 6 (a) to (k). The frequency error between the two methods is shown in Fig. 7. The results agree well, with a maximum frequency error of 0.06Hz. The displacement time history measured by the proposed system is compared with that measured by the accelerometer, and this was shown in Fig. 8. The results correlate well, with an average amplitude error of 0.012 mm. It is hence concluded that, the proposed system can accurately measure such vibration within the frequency range 1 Hz to 150Hz.
Fig. 6 Comparison of frequency measurement results of vibration measurement accuracy and capacity verification experiment: (a) about 1Hz, (b) about 4Hz, (c) about 7Hz, (d) about 10Hz, (e) about 30Hz, (f) about 50Hz, (g) about 70Hz, (h) about 90Hz, (i) about 110Hz, (j) about 130Hz, (k) about 150Hz.
This vibration measurement accuracy and capacity verification experiment demonstrates that the proposed system can measure dynamic vibration with good accuracy. Nevertheless, because the improved LKBOPT is a sub-pixel optical flow estimation method, its accuracy is related to the pixel resolution but not limited by it.

3.3 Simultaneous multipoint vibration measurement experiment of a typical TWPs

The vibration measurement effectiveness and capacity for TWPs of the proposed system is then evaluated through a simultaneous multipoint vibration measurement experiment of an aircraft T tailplane model. An aircraft T tailplane model is built with stainless steel plate of 1mm thickness. The proposed system is experimentally evaluated against the traditional accelerometer sensor in this experiment too. The vibration can be recorded simultaneously by the accelerometer sensor and the high-speed camera system for comparison. The shaker with aircraft T tailplane is shown in Fig. 9(a). Two accelerometer sensors are mounted on both sides of the T plane model, and the whole model is fixed on the shaker. The shaker is driven by the signal generator. Then the accelerometer vibration signal is sampled at 375Hz with the dynamic data acquisition system and streamed into the computer through the USB cable and recorded by the general control & analysis software supplied with the system. Meanwhile, the vibration of the whole model is also recorded with the high-speed camera. The video images captured by the camera are digitized into 640×480 pixel images in 8-bit grey scales at 375 fps and streamed into the computer through an USB 3.0 cable. The whole experimental rig is shown in Fig. 9(b). The sampling time of both two are set as 16 second. The motion information of the accelerometer sensor is extracted by performing the proposed system on the connected computer. A typical image captured by the camera is displayed in Fig. 9(c), in which the red box is selected as the template and the green points are selected as the tracking features.
The vibration frequency of the shaker is set as about 30.9Hz. And then the acceleration signals are transferred into displacement signal for comparison too. Then both of the displacement signals are processed with FFT. The displacement time history measured by the proposed system are compared with those measured by the accelerometer sensors, which is shown in Fig. 10(a) and (b), the amplitude frequency spectrum is compared and shown in Fig. 10(c) and (d). The results agree well, and the frequency error is 0.04 Hz and the average displacement error is 0.025 mm. Therefore, the proposed system could accurately measure multipoint vibration of TWPs simultaneously, having demonstrated its good measurement capacity.

4. Conclusion

In this study, in order to satisfy the simultaneous multipoint vibration measurement requirements of TWPs, a coarse-to-fine tracking method is developed by introducing LKBOPT to vibration measurement and
combining it with LKBTT. A region of interest detection algorithm is investigated to acquire the tracking template according to the measurable object. LKBTT is introduced to realize the coarse tracking and obtain the vibration information of the tracking template. An improved LKBOPT is investigated to realize the fine tracking and acquire the vibration information of multipoint inside the tracking template simultaneously. Finally, the simultaneous measurement system of multipoint mechanical vibration for TWPs is utilized with a high-speed camera system based on both Python and open source computer vision (OpenCV) libraries. The proposed method has some advantages. First of all, the proposed method can realize simultaneous multipoint vibration measurements of TWPs. Secondly, the size of the template is decided by the detected interesting object which provide contextual cues for optical flow estimation. Finally, the motion information of the object of interesting or specific multipoint inside the template can be extracted with proposed method.

Three experimental studies validated the performance of the proposed system. The vibration measurement accuracy and capacity of the proposed system was first evaluated through a shaker vibration measurement experiment. The vibration measurement effectiveness and capacity for TWPs of the proposed system is then evaluated through a simultaneous multipoint vibration measurement experiment of an aircraft T tailplane model. The results using the proposed system achieved excellent agreement with those of traditional accelerometer sensors.

The changes in shading, lighting, and background conditions in the field made it difficult to extract accurate vibration from video images using the proposed system, further work will be carried out to improve the robustness of the proposed method to brightness variations.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data Availability Statement

The data used to support the findings of this study were supplied by Jigang Wu under license and so cannot be made freely available. Requests for access to these data should be made to Jigang Wu by jwu@cvm.ac.cn.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (NSFC) (Grant no. 51775181).

References

[1] X. Wang, Q. Song and Z. Liu. “Dynamic model and stability prediction of thin-walled component milling with multi-modes coupling effect,” Journal of Materials Processing Technology, vol. 288, pp. 116869, 2021.
[2] G. Wang, W. Li, G. Tong and C. Pang. “Improving the machining accuracy of thin-walled parts by online measuring and allowance compensation,” The International Journal of Advanced Manufacturing Technology, vol. 92, no. 5, pp. 2755-2763, 2017.
[3] N. Tandon and A. Choudhury. “A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings,” Tribology International, vol. 32, no. 8, pp.469-480, 1999.
[4] J. Liu and X. Yang. “Learning to see the vibration: a neural network for vibration frequency prediction,” Sensors, vol. 18, no. 8, pp. 2530, 2018.
[5] J. Abir, S. Longo, P. Morantz and P. Shore. “Optimized estimator for real-time dynamic displacement measurement using accelerometers,” Mechatronics, vol. 39, pp. 1-11, 2016.
[6] C. Warren, C. Niezrecki, P. Avitabile and P. Pingle. “Comparison of FRF measurements and mode shapes determined using optically image based, laser, and accelerometer measurements,” Mechanical Systems and Signal Processing,
vol. 25, no. 6, pp. 2191-2202, 2011.

[7] H. Kayaba and Y. Kokumai. “Non-contact full field vibration measurement based on phase-shifting,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3655-3663, 2017.

[8] J. J. Lee and M. Shinozuka. “Real-time displacement measurement of a flexible bridge using digital image processing techniques,” Experimental Mechanics, vol. 46, no. 1, pp. 105-114, 2006.

[9] D. Feng, M. Q. Feng, E. Ozer and Y. Fukuda. “A vision-based sensor for noncontact structural displacement measurement,” Sensors, vol. 15, no. 7, pp. 16557-16575, 2015.

[10] D. L. B. R. Jurjo, C. Magluta, N. Roitman and P. B. Goncalves. “Analysis of the structural behavior of a membrane using digital image processing,” Mechanical Systems and Signal Processing, vol. 54-55, pp. 394-404, 2015.

[11] K. S. Son, H. S. Jeon, G. S. Chae, J. S. Park and S. O. Kim. “A fast high-resolution vibration measurement method based on vision technology for structures,” Nuclear Engineering and Technology, vol. 53, no. 1, pp. 294-303, 2021.

[12] H. V. de Figueiredo, D. F. Castillo-Zúñiga, N. C. Costa, O. Saotome and R. G. A. da Silva. “Aeroelastic vibration measurement based on laser and computer vision technique,” Experimental Techniques, vol. 45, pp. 95-107, 2021.

[13] Z. Qiu, X. Wang, X. M. Zhang and J. Liu. “A novel vibration measurement and active control method for a hinged flexible two-connected piezoelectric plate,” Mechanical Systems and Signal Processing, vol. 107, pp. 357-395, 2018.

[14] G. Wu, X. Xiao and C. Ma. “Vision-based 2D Vibration Displacement Measurement of Hoisting Vertical Rope in Mine Hoist,” Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition, pp. 210-214, 2020.

[15] M. A. Kuddus, J. Li, H. Hao, C. Li and K. Bi. “Target-free vision-based technique for vibration measurements of structures subjected to out-of-plane movements,” Engineering Structures, vol. 190, pp. 210-222, 2019.

[16] B. K. P. Horn and B. G. Schunck. “Determining optical flow,” Artificial Intelligence, vol. 17, no. 1-3, pp. 185-203, 1981.

[17] B. D. Lucas and T. Kanade. “An iterative image registration technique with an application to stereo vision,” Proceedings DARPA Image Understanding Workshop, pp. 121-130, 1981.

[18] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black and R. Szeliski. “A database and evaluation methodology for optical flow,” International Journal of Computer Vision, vol. 92, no. 1, pp. 1-31, 2011.

[19] S. Baker and I. Matthews. “Lucas-kanade 20 years on: A unifying framework,” International Journal of Computer Vision, vol. 56, no. 3, pp. 221-255, 2004.

[20] J. Morllet and G. Michon. “Virtual vibration measurement using KLT motion tracking algorithm,” Journal of Dynamic Systems, Measurement, and Control, vol. 132, no. 1, pp. 1-8, 2010.

[21] X. Lei, Y. Jin, J. Guo and C. Zhu. “Vibration extraction based on fast NCC algorithm and high-speed camera,” Applied Optics, vol. 54, no. 27, pp. 8198-8206, 2015.

[22] J. Guo and C. Zhu. “Dynamic displacement measurement of large-scale structures based on the Lucas–Kanade template tracking algorithm,” Mechanical Systems and Signal Processing, vol. 66-67, pp. 425-436, 2016.

[23] D. Zhang, J. Guo, X. Lei and C. Zhu. “A high-speed vision-based sensor for dynamic vibration analysis using fast motion extraction algorithms,” Sensors, vol. 16, no. 4, pp. 572-589, 2016.

[24] C. Zhu, J. Guo, D. Zhang, Y. Shen and D. Liu. “In situ measurement of wind-induced pulse response of sound barrier based on high-speed imaging technology,” Mathematical Problems in Engineering, vol. 2016, pp. 1-8, 2016.

[25] D. Feng and M. Q. Feng. “Vision-based multipoint displacement measurement for structural health monitoring,” Structural Control and Health Monitoring, vol. 23, no. 5, pp. 876-890, 2016.

[26] Y. Yang and X. B. Yu. “Image analyses for video-based remote structure vibration monitoring system,” Frontiers of Structural and Civil Engineering, vol. 10, no. 1, pp. 12-21, 2016.

[27] T. Brox and J. Malik. “Large displacement optical flow: descriptor matching in variational motion estimation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 3, pp. 500-513, 2011.

[28] J. Y. Bouguet. “Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm,” Intel Corporation, vol. 5, pp. 1-10, 2001.