A Vision-Based Vibration Measurement Method for Large Hydraulic Manipulator

Shuwei Yang, Weidi Huang*, Junhui Zhang, Bing Xu, Yaoxing Wang, and Qi Wang

State Key Laboratory of Fluid Power and Mechatronic System, Zhejiang University, Hangzhou, China
E-mail: swyang@zju.edu.cn, wdhuang@zju.edu.cn, benzjh@zju.edu.cn, bxu@zju.edu.cn, zjuyxwang@zju.edu.cn, qiwang98@zju.edu.cn.

Abstract

Vibration frequently occurs at the Tool Center Point (TCP) of large hydraulic manipulators during operation, making the monitoring of TCP vibrations imperative for ensuring the safety and operational efficiency of the manipulator. Traditional methods for TCP vibration measurement generally rely on contact-based techniques, such as Inertial Measurement Units (IMU). However, these methods face limitations, including an inability to directly quantify vibration amplitudes and susceptibility to noise interference, which can negatively impact the manipulator's performance. To address these challenges, the present study proposes a novel, non-contact vibration measurement technique that utilizes vision-based technology. Specifically, a high-resolution camera is employed to capture images of the manipulator's TCP, with the Harris corner detection algorithm applied to identify key features. The changes in these features are then tracked and converted into positional shifts within the manipulator's coordinate system. This approach ensures the precision and stability of the measurement and prevents any direct interference with the equipment's operation. Experimental results across various operational scenarios reveal that, when compared to conventional methods, the proposed vision-based technique offers superior accuracy and stability in non-contact settings. Additionally, it demonstrates enhanced performance in vibration data extraction and processing efficiency. The method delineated in this study provides a viable solution for the vibration measurement and maintenance of large hydraulic manipulators, offering novel insights into enhancing operational safety and precision control.

Keywords: Vibration measurement, Image recognition, Manipulator kinematics, Signal processing.

1 Introduction

Large hydraulic manipulators are widely used in industries such as material handling, port loading and unloading, and municipal construction, often operating in high-load environments [1,2]. As a result, ensuring the stability and safety of these manipulators has become a critical concern. The nonlinear characteristics of the hydraulic system, combined with the coupling effects of the multi-link structure, lead to low-frequency vibrations (typically between 2-25Hz) during transient operations such as start-up, stop, and reversal. Although these vibrations generally have minimal impact on the manipulator's overall functionality, they can cause issues such as reduced operational accuracy, structural fatigue, and hydraulic pipeline leaks. Therefore, the development of high-precision vibration measurement methods, suitable for complex industrial environments, has become an urgent technical challenge that must be addressed.

Traditional vibration monitoring techniques commonly employ contact sensors, such as piezoelectric accelerometers and fiber optic IMUs, for vibration measurement [3,4]. While these methods offer certain advantages in low-frequency vibration monitoring, their inherent limitations restrict their widespread application in complex industrial environments. Contact sensors generally require precise mounting on the manipulator's end effector, a process that demands high accuracy during installation. In specific operational environments, such as those with elevated temperatures, oil contamination, and intense vibrations, these sensors are particularly susceptible to interference and signal drift. Additionally, the installation process may alter the mass distribution of the measured object, introducing vibrational disturbances that compromise measurement accuracy. Conventional contact methods, like IMUs, provide only acceleration data, necessitating secondary integration to estimate displacement amplitudes. This process is prone to cumulative errors and cannot directly yield the actual vibration amplitude. In large workspace environments, the installation and wiring process is often complex, requiring cables to be routed along the entire length of the manipulator. This process is both time-consuming and labor-intensive, and it may disrupt the normal operation of the manipulator.

Non-contact devices, such as the Laser Doppler Vibrometer (LDV), offer micron-level measurement accuracy and are capable of effectively monitoring vibrations in complex environments [5]. However, the high cost of these devices (typically exceeding $100,000 per unit), intricate optical calibration procedures, and sensitivity to line-of-sight obstructions present significant challenges for their widespread use in practical construction scenarios. As a result, the cost-performance ratio of these devices is substantially lower than that of traditional contact sensors.

In recent years, with the rapid development of computer vision technology, vision-based vibration measurement methods have emerged as a significant research direction [6-8]. These methods extract and analyze the motion trajectories of target feature points in image sequences to directly obtain spatial displacement information. They offer notable advantages, such as non-contact measurement, full-field monitoring, and low cost. Vision-based vibration measurement methods can currently be categorized into two types: the marker-based template matching method, which relies on manually calibrated markers and is susceptible to occlusion in complex scenarios [9], and the markerless tracking method, which utilizes optical flow. The latter is independent of manual intervention and offers enhanced flexibility [10]. However, it is vulnerable to noise interference when employing the traditional Lucas-Kanade algorithm, leading to tracking failures in scenarios with significant displacements [11,12]. The operation of the hydraulic manipulator typically results in low-frequency, small-amplitude oscillations at the TCP, which imposes stringent requirements on the positioning accuracy and real-time performance of vision algorithms.

To overcome the limitations of existing methods, this paper proposes a vision-based vibration measurement method integrating Harris corner detection with optical flow tracking. The Harris corner detection algorithm is used to extract target feature points, and optical flow tracking enables precise tracking, allowing for efficient and accurate extraction of vibration displacement data. Unlike traditional contact sensors, the vision-based measurement method does not require direct contact with the equipment or environment, enabling stable operation in harsh conditions with strong robustness. To validate the accuracy of the method, a comparison with existing contact-based measurement methods is conducted. Experimental results demonstrate that the vision-based vibration measurement method ensures accuracy and stability while exhibiting excellent performance in vibration data extraction and processing efficiency. This approach offers an innovative solution for vibration measurement and maintenance of large hydraulic manipulators, providing effective technical support for enhancing the safety and precision control of equipment.

The structure of this paper is as follows: Chapter 2 presents the research object and technical approach, providing a detailed explanation of the proposed method. Chapter 3 outlines the specific implementation steps for vision-based vibration measurement. Chapter 4 describes the procedure for contact-based vibration measurement. Chapter 5 details experiments conducted under various operating conditions to validate the effectiveness of the proposed method. Chapter 6 concludes the paper and discusses future research directions.

2 Technical approach

The proposed methodology employs a vision-based, non-contact vibration measurement technique for high-precision monitoring of the TCP of large hydraulic manipulators (Figure 1). A camera is used to capture real-time images of the manipulator's end effector, and image processing techniques are applied to detect small vibration variations. The integration of Harris corner detection and optical flow-based feature tracking allows for the extraction of vibration displacement data from the TCP without requiring direct contact with the equipment, providing an efficient and stable solution for vibration measurement.

Figure 1: Concrete pumping equipment fitted with a large hydraulic manipulator.

Initially, markers with distinct geometric features are affixed to the TCP of the manipulator to serve as target points for subsequent image processing. A high-resolution camera is used during the image acquisition process to ensure an adequate frame rate and resolution, allowing for precise capture of small vibrations at the TCP. Each image frame is processed using the Harris corner detection algorithm to extract feature points, which correspond to the positions of the markers on the manipulator's TCP. This process forms the foundation for subsequent displacement calculations.

The Lucas-Kanade algorithm, based on optical flow, is then applied to track the extracted feature points and capture their motion trajectories across adjacent image frames [13,14]. Optical flow operates under the assumption of image brightness consistency and calculates pixel displacements between consecutive frames to derive relative position changes of the TCP. Given that the vibrations of hydraulic manipulators typically manifest as minute displacements, optical flow can accurately track these subtle movements, ensuring high precision in vibration measurement.

This study utilizes camera calibration to obtain the camera's intrinsic and extrinsic parameters, employing geometric transformation methods to convert displacement information from the image coordinate system into the manipulator's work coordinate system. The camera's intrinsic parameters (e.g., focal length and principal point) and extrinsic parameters (describing the relationship between the camera and the world coordinate systems) are used to achieve precise mapping from the image plane to the actual working space through perspective transformation and other algorithms.

To validate the accuracy of the proposed method, a comparison is made between the vibration measurement results obtained using the vision-based method and the experimental results from a contact-based measurement method (IMU). As a traditional vibration measurement tool, the IMU provides dynamic information, such as acceleration. However, due to the error accumulation characteristics and the inherent limitations of the acceleration integration method, significant errors may be introduced over long measurement durations. To address this, signal processing techniques are applied to extract vibration components from the IMU data. By combining the manipulator's kinematic model with the angular data provided by the joint sensors, displacement data for the TCP is obtained, excluding vibration signals. The comparison of experimental results from both the vision-based and contact-based methods confirms the superiority of the former in terms of accuracy and effectiveness.

3 Vision-based measurement method

Vibration measurement of the TCP requires non-contact vibration signal detection using computer vision technology, which involves precise camera calibration and image processing techniques to extract displacement information. This section provides a detailed explanation of the camera calibration methods, including intrinsic and extrinsic parameters, Harris corner detection principles, and optical flow applications. Additionally, the relationship between the checkerboard markers and the camera matrix is utilized to convert pixel coordinates into the manipulator's working coordinate system, enabling precise measurement of TCP position variations.

3.1 Camera intrinsic and extrinsic parameters calibration

To accurately convert pixel information from the image coordinate system into physical coordinates, the intrinsic and extrinsic parameters of the camera must first be calibrated. The intrinsic parameters include the focal length, principal point, and distortion coefficients. The extrinsic parameters describe the relationship between the camera coordinate system and the world coordinate system. Calibration of both intrinsic and extrinsic parameters can be achieved using a calibration board with known dimensions, with the Zhang checkerboard calibration method [15] being a widely used approach.

During the calibration process, a calibration board with a known square size is used, and multiple images are captured to extract the positions of the chessboard corners. The 2D coordinates of the chessboard corners in the image are denoted as (u_i, v_i), and the corresponding 3D coordinates in the world coordinate system are (X_i, Y_i, Z_i). The intrinsic and extrinsic parameters of the camera satisfy the following projection relationship:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \tag{1}$$

Where $\boldsymbol{K}$ is the intrinsic parameter matrix of the camera, represented as:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Where f_x and f_y represent the focal lengths, and c_x and c_y represent the principal point coordinates. $\boldsymbol{R}$ and $\boldsymbol{t}$ are the rotation matrix and translation vector in the camera’s extrinsic parameters, respectively, which describe the transformation from the camera coordinate system to the world coordinate system.

By minimizing the reprojection error, the intrinsic and extrinsic parameters of the camera are computed, enabling the conversion from image coordinates to world coordinates.

3.2 Harris corner detection and optical flow tracking

The Harris corner detection algorithm used in this paper is a widely applied feature point extraction method in image processing. Corners are detected by computing the local autocorrelation matrix of each pixel in the image. The image gradients in the x and y directions are represented by Ix and Iy, respectively. The autocorrelation matrix within the image window is expressed as:

$$M = \begin{bmatrix} \Sigma I_x^2 & \Sigma I_x I_y \\ \Sigma I_x I_y & \Sigma I_y^2 \end{bmatrix} \tag{3}$$

Where $\sum I_x^2$ is the sum of the squared gradients in the x-direction within the window, and $\sum I_x I_y$ is the sum of the product of the gradients in the x and y directions. By calculating the eigenvalues of the matrix M, it is possible to determine whether a point is a corner. Specifically, the Harris corner response function is defined as:

$$R = \det(M) - k(\mathrm{trace}(M))^2 \tag{4}$$

Where k is an empirical constant, typically in the range of 0.04 to 0.06, det(M) is the determinant of the matrix M, and trace(M) is the trace of matrix M. In the event that the value of R is sufficiently large, the point is designated as a corner, thereby indicating the presence of significant features in the image. In the context of vibration measurement, the markers situated at the TCP of the manipulator can be utilized to extract multiple stable corner points through the implementation of Harris corner detection. These corner points subsequently serve as feature points for the purpose of subsequent optical flow tracking, thereby providing the foundational data necessary for the calculation of TCP position changes.

After corner points are detected in each image frame, optical flow is applied to compute the pixel displacements between adjacent frames, thereby tracking the movement of objects in the image. The fundamental assumption of optical flow is that the intensity values between adjacent frames remain constant, meaning the object in the image undergoes translational motion within a short time. Based on this assumption, the intensity values of a feature point in the image at times t and $t+1$ are represented by $I(x, y, t)$ and $I(x+\Delta x, y+\Delta y, t+1)$, respectively. The fundamental assumption of optical flow is as follows:

$$I(x,y,t) = I(x+\Delta x, y+\Delta y, t+1) \tag{5}$$

Performing a Taylor expansion and neglecting higher-order terms results in the optical flow equation:

$$I_x u + I_y v + I_t = 0 \tag{6}$$

Where u and v represent the velocities of the point in the x and y directions, respectively, and I_x and I_y are the image gradients in the x and y directions, while I_t is the time gradient. By using the pixel gradients within a local window, optical flow can be applied to calculate the displacement of each feature point in the image, thus providing the vibration displacement information of the TCP.

3.3 Coordinate transformation

The integration of Harris corner detection and optical flow tracking facilitates the acquisition of variations in the coordinates of the corner points in the image plane. These corner points represent changes in the position of the TCP of the manipulator, and by analyzing the trajectory of these points, the displacement of the TCP can be accurately tracked. In this paper, the intrinsic and extrinsic parameter matrices derived from the camera calibration process are used to convert pixel coordinates in the image plane to 3D physical coordinates in the world coordinate system. These parameter matrices describe the relationship between the camera coordinate system and the world coordinate system, enabling effective mapping of pixel points in the image coordinate system to the physical workspace. A detailed explanation of the coordinate transformation process is provided in Equations 1 and 2.

4 Contact Measurement method

When using the IMU for measurements, it must be securely mounted on the target to ensure it moves in conjunction with the object being measured [16]. The IMU is capable of recognizing acceleration signals that encompass the rigid motion of the arm, external disturbances, and vibration components. These signals are highly susceptible to low-frequency drift (e.g., static acceleration caused by gravity) and high-frequency noise. To mitigate the impact of these factors on the measurement results, the acquired acceleration signals are preprocessed, with only the vibration-related valid signals being extracted. The IMU is positioned near the end of the manipulator to capture its acceleration signals. Through data processing, the vibration of the Tool Center Point (TCP) is characterized. Angle encoders are installed at the joints of the manipulator to measure displacement signals that do not include vibration components. The rigid displacement of the arm is determined through a combination of joint angle information and the kinematic model. This dual approach is integrated to ensure the precision of the TCP displacement.

4.1 Bandpass filtering

Band-pass filtering serves as the initial phase in the processing of IMU acceleration signals. Since the vibrations of the manipulator typically occur within a specific frequency range, the band-pass filter is used to efficiently eliminate undesired low-frequency and high-frequency components. Low-frequency components may include gravitational acceleration or slower rigid body motion signals, while high-frequency components may arise from noise or other disturbances.

In this study, a Butterworth band-pass filter is applied due to its smooth and flat frequency response within the passband, which ensures that only the relevant vibration frequencies are retained while effectively suppressing unwanted components. This configuration helps eliminate low-frequency components associated with the manipulator's rigid motion and sensor noise, thereby preserving the frequency components related to vibration. The band-pass filtered acceleration signal primarily contains components that correspond to the vibration frequency of the manipulator, providing a clear signal foundation for subsequent displacement extraction.

4.2 Freuqency-domain processing and double integration

Given that acceleration is defined as the second derivative of displacement, direct integration of the acceleration signal in the time domain introduces drift errors, which can lead to inaccurate results. To address this issue, a frequency-domain integration method for displacement extraction is proposed in this paper. This approach involves transforming the band-pass filtered acceleration signal into the frequency domain using the Fourier transform, which converts the time-domain signal into a frequency-domain signal, allowing for more precise manipulation of each frequency component. In the frequency domain, the acceleration signal $A(f)$ is represented as a summation of frequency components, where f denotes the frequency.

The primary advantage of frequency-domain analysis is that it avoids the error accumulation that occurs during time-domain integration. In the frequency domain, integration is equivalent to multiplying each frequency component, a process that is more stable and accurate. This method effectively eliminates the drift errors introduced during time-domain integration.

In the frequency domain, each frequency component of the acceleration signal is integrated. Since acceleration is the second derivative of displacement, two integrations are required. Initially, the processed acceleration signal $A(f)$ is integrated to yield the velocity signal $V(f)$:

$$V(f) = \frac{A(f)}{i2\pi f} \tag{7}$$

The velocity signal $V(f)$ is then integrated again to obtain the displacement signal $D(f)$:

$$D(f) = \frac{V(f)}{i2\pi f} = \frac{A(f)}{(i2\pi f)^2} \tag{8}$$

Once the vibration displacement signal D(f) has been obtained in the frequency domain, the inverse Fourier transform is applied to convert the frequency-domain signal back to the time domain, yielding the final vibration displacement signal. The inverse Fourier transform recovers the individual frequency components in the frequency domain and reconstructs a continuous signal in the time domain. The resulting time-domain signal represents the pure vibration displacement, effectively eliminating errors caused by initial conditions and noise. The frequency-domain integration method facilitates the precise extraction of vibration components from the acceleration signal, bypassing the error accumulation that can occur with time-domain integration.

4.3 Rigid displacement calculation

During the movement of the manipulator, the positional changes of the TCP involve not only the minor variations caused by vibrations but also substantial changes resulting from rigid body motion. Therefore, after obtaining the pure vibration displacement signal from the IMU data, it is necessary to calculate the rigid motion changes of the manipulator based on the angular information between the joints (i.e., positional changes that exclude vibration components).

In this study, a kinematic model is developed for the mechanical structure of the large manipulator (Figure 2). The three sections of the manipulator's arms can rotate around the joints, and hydraulic actuators drive the amplitude-changing mechanism, which controls the lifting and lowering of each arm section. The joints are subject to angular constraints, and by utilizing the angular information between the joints, the TCP position can be computed by combining the arm lengths, as expressed by the following equation:

$$\begin{cases} x_{TCP} = c_0(a_3c_{123} + a_2c_{12} + a_1c_1 + a_0) \\ y_{TCP} = s_0(a_3c_{123} + a_2c_{12} + a_1c_1 + a_0) \\ z_{TCP} = a_3s_{123} + a_2s_{12} + a_1s_1 + d_0 \\ \xi = \varphi_1 + \varphi_2 + \varphi_3 \end{cases} \tag{9}$$

where φ_0, φ_1, φ_2, φ_3 denote angular vectors of joints 1-3, $\boldsymbol{P} = [x, y, z]^{\mathrm{T}} \in \mathbf{R}^3$ specifies the TCP position. $\zeta \in (-\pi/2,\ \pi/2)$ represents the inclination angle between the terminal arm section and horizontal plane. The specific corresponding parameters are shown in the table. φ_{lim} define rotational constraints per joint. All physical parameters are quantified in Table 1.

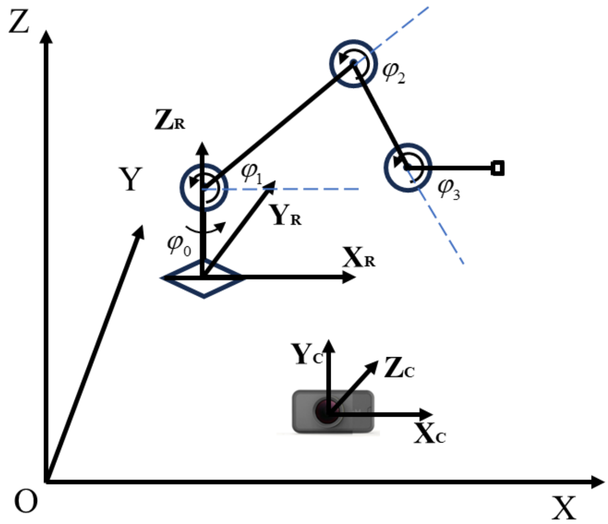


Figure 2: D-H coordinates of manipulator.

Table 1: D-H parameters.

Link	$\varphi_i(°)$	$d_i(mm)$	$a_i(mm)$	$\alpha_i(°)$	$\varphi_{lim}(°)$
0	φ_0	2100	0	90	-90 ~ 90
1	φ_1	0	5500	0	0 ~ 88
2	φ_2	0	3750	0	-180 ~ 0
3	φ_3	0	3700	0	-60 ~ 180

The rigid displacement, obtained from the forward solution of the manipulator's kinematics, and the vibration displacement, obtained from the frequency-domain integration of the IMU data, are accumulated to calculate the TCP position change, including the vibration displacement.

5 Experiment

5.1 Experimental setup

The experimental apparatus in this study consists of a concrete pumping system equipped with a manipulator. The device features a three-degree-of-freedom rotational manipulator, with a total length of 13.5 meters when fully extended. Each section of the manipulator is elevated and lowered by hydraulic actuators, combined with an amplitude-changing mechanism. The rotational motion of the manipulator is driven by a rotary hydraulic motor that operates a gear system. The technical specifications of the manipulator are detailed in Table 2. To obtain angular information between the joints, angle encoders are installed, and an IMU sensor is positioned near the TCP of the manipulator. The experimental setup and the installation of the corresponding sensors are shown in Figure 3.

The controller sampling frequency is 1000 Hz, the IMU sampling frequency is 500 Hz, and the camera sampling frequency is 60 Hz. The sampling frequency of the control multiway valves is 20 Hz.

Table 2: Experimental platform structure parameters.

Nomenclature	Value
Boom 1 length l_1	5.55 m
Boom 2 length l_2	3.75 m
Boom 3 length l_3	3.70 m
Teeth of Rotary Hydraulic Motor Gear z_1	15
Teeth of Slewing Platform Gear z_2	124
Rotary Mechanism Ratio	15/124

To validate the effectiveness of the vision-based vibration measurement method, two sets of experimental conditions are conducted. For each set, vibration measurements are performed using both vision-based and contact-based methods, and the results from both methods are compared and validated. The experimental conditions are as follows:

Experiment A: The camera is placed on the side of the manipulator, with a vertical distance between the manipulator and the camera ranging from 5m to 6m. This distance range simulates a typical scenario where the camera is closer to the manipulator, allowing it to capture higher-resolution images and clearer vibration information. This experiment is designed to verify that the vision-based vibration measurement method can achieve higher accuracy within the ideal distance range and provide an effective comparison with the contact method.

Experiment B: The camera is placed on the side of the manipulator, with a vertical distance between the manipulator and the camera exceeding 10m. This configuration simulates a scenario where vibration measurements are taken at longer distances. This setup allows for the evaluation of the vision-based approach at greater distances, along with the challenges it may face in real-world applications, such as image blur and degradation of vibration tracking accuracy.

Prior to the commencement of the experiments, the intrinsic and extrinsic parameters of the camera were calibrated. The intrinsic parameter matrix was obtained by capturing images of the checkerboard pattern in various poses and employing the Zhang calibration method. During the calibration process, the coordinates of multiple corner points on the checkerboard were measured, providing intrinsic parameters such as focal length and principal point position. The mean error in pixels is shown in Figure 4. The extrinsic parameter matrix was determined based on the relative position between the camera and the manipulator. This process ensured the accuracy and reliability of the camera measurements. To ensure precision, checkerboard markers were affixed near the IMU installation point at the TCP. During the experiment, the Harris corner detection algorithm was used to detect the four corner points of the black square at the center of the checkerboard, and optical flow was applied to track the motion trajectories of these corner points in real-time.

Considering the IMU sensor errors and noise, the IMU was initialized before its signals were used to calculate the TCP vibration displacement. Acceleration signals from the IMU were collected while the manipulator was stationary to eliminate errors caused by static gravity or other factors. During the experiment, acceleration signals were collected in real-time, and in the subsequent offline processing stage, the data were filtered and integrated. A band-pass filter was applied to remove low-frequency static acceleration and high-frequency noise, ensuring the signal's accuracy. Subsequently, the frequency-domain integration method was employed to convert the acceleration signal into displacement data, thereby obtaining the pure vibration displacement signal.

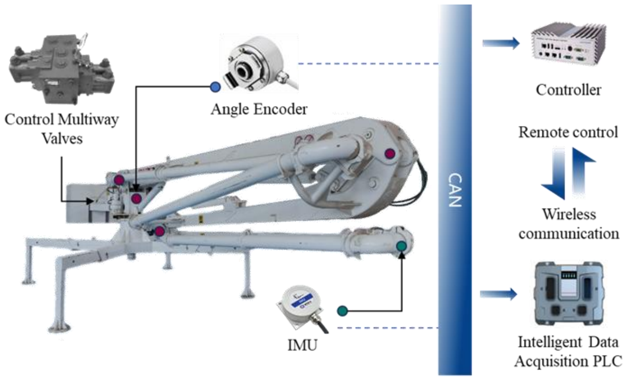
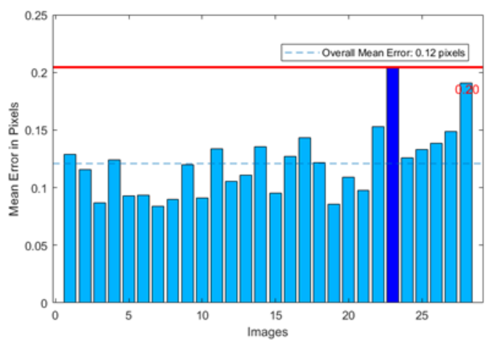
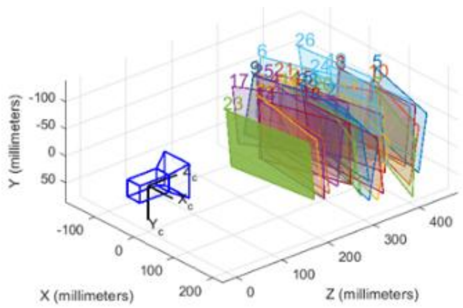


Figure 3: Experimental platform setup on the concrete pumping equipment.



(a) Camera inherent parameter calibration error



(b) Schematic diagram of camera parameter calibration

Figure 4: Camera inherent parameter calibration and parameter error.

5.2 Experiment A

During the motion of the manipulator, the vibration amplitude in the horizontal direction is found to be less significant than in the vertical direction. To provide a clear comparison of the effectiveness of the vision-based and contact-based measurement methods, the present experiment focuses on the extraction of the more prominent vertical vibrations. The results of Experiment A are shown in Figures 5-9.

In Experiment A, the motion duration of the manipulator is 50 seconds. During the start-up and stop phases, slight vibrations were observed, even when minimal TCP velocity was set, due to the impact of the hydraulic system. These vibrations are evident in the acceleration signal shown in Figure 6, predominantly occurring between 0-10 s and 40-50 s. Figure 7 illustrates the Z-direction vibration displacement changes obtained after processing the acceleration signal collected by the IMU. Figure 8 displays the angular data collected by the joint angle encoders of the manipulator. Figure 9 presents a comparison of the measurement results obtained from both the vision-based method and the contact-based methods (IMU and angle encoders).

The experimental results demonstrate that the proposed method exhibits nearly identical overall trends in the experimental curve when compared to the contact-based measurement methods. During the initial and final stages of the experiment, slight vibrations were observed in the TCP of the manipulator. The results indicate that both the vision-based and contact-based measurement methods accurately represented the vibration trends and amplitudes.

Errors in the vibration amplitudes measured by both methods were observed during certain time periods. This phenomenon is primarily attributed to the accumulation of errors in the IMU during extended measurement periods. As a result, the amplitude may be either overestimated or underestimated.

5.3 Experiment B

The specific experimental process for Experiment B is illustrated in Figure 10, with the manipulator's motion duration set to 60 seconds. The vibration of the manipulator's TCP is evident in the acceleration signal shown in Figure 11, predominantly occurring between 10-20 s and 50-60 s. Figure 12 presents the Z-direction vibration displacement changes derived from the processed acceleration signal collected by the IMU. The angular data collected by the joint angle encoders of the manipulator is shown in Figure 13. Figure 14 presents the measurement results obtained from both the vision-based and contact-based methods (IMU and angle encoders), along with a comparative analysis of the results.

Similar to Experiment A, slight vibrations were observed at the TCP of the manipulator during the initial and final stages of the experiment. The results indicate that both the vision-based and contact-based measurement methods accurately represented the vibration trends and amplitudes. However, some errors in the vibration amplitude were still observed during certain periods. This discrepancy is primarily attributed to the accumulation of errors within the IMU during prolonged measurement periods, resulting in the overestimation or underestimation of the amplitude at specific instances.

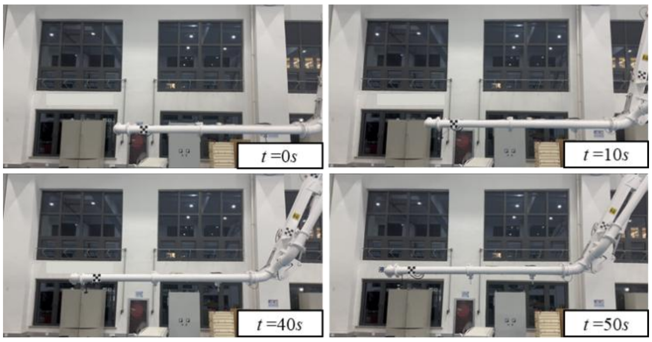


Figure 5: Manipulator motion process.

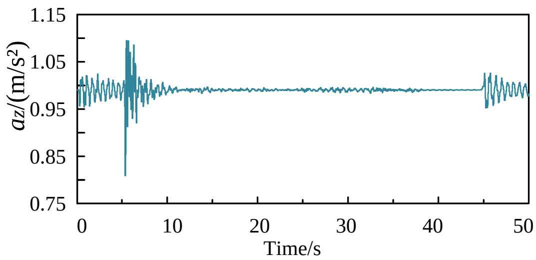


Figure 6: Z-direction acceleration variation.

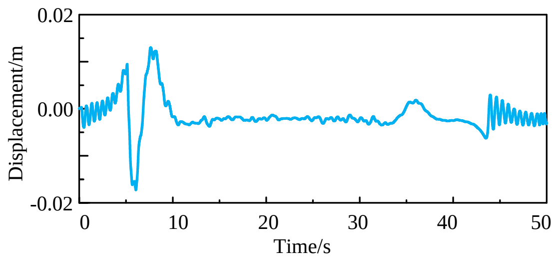


Figure 7: Variation of vibration displacement.

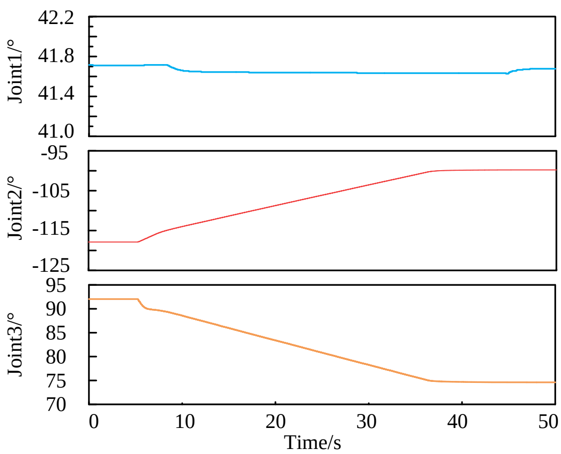


Figure 8: Manipulator Joint Angle Variation.

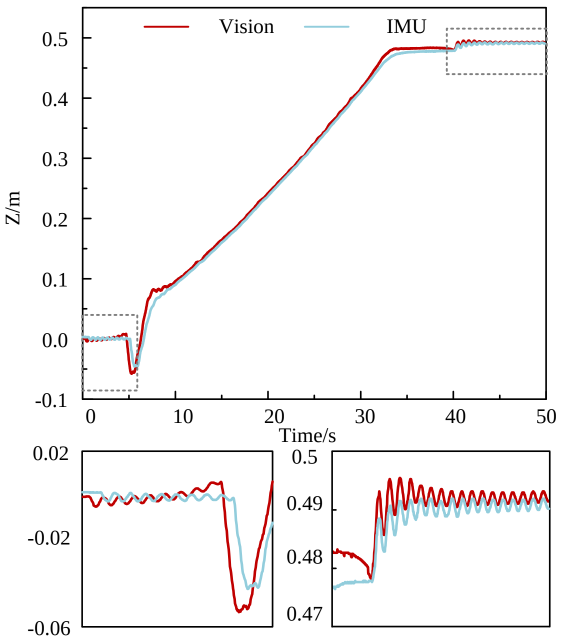


Figure 9: Position of the manipulator's TCP in the Z-direction.

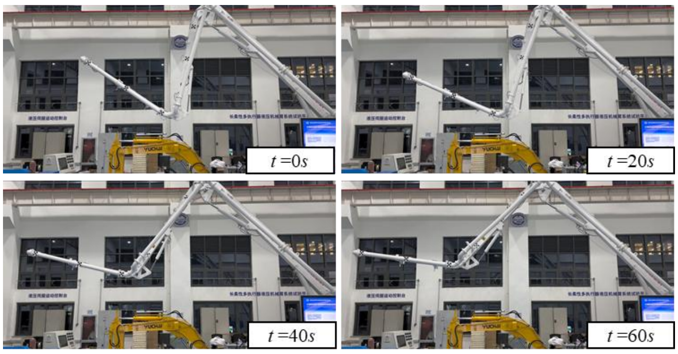


Figure 10: Manipulator motion process.

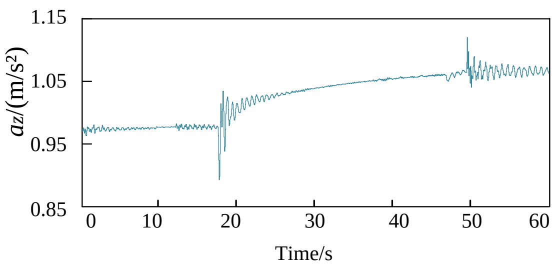


Figure 11: Z-direction acceleration variation.

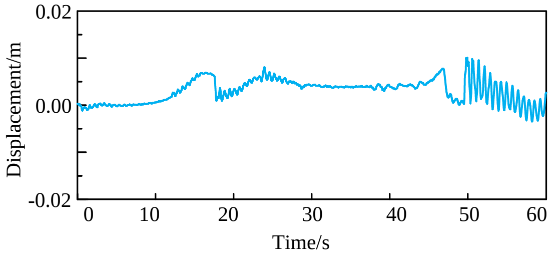


Figure 12: Variation of vibration displacement.

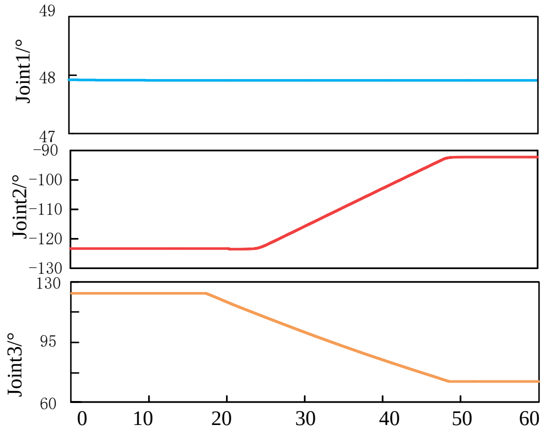


Figure 13: Manipulator Joint Angle Variation.

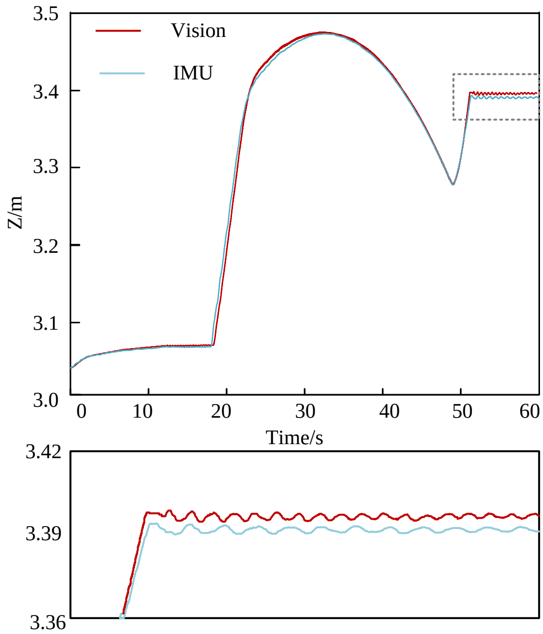


Figure 14: Position of the manipulator's TCP in the Z-direction.

5.4 Experiment

The experimental data demonstrate that both the vision-based and IMU methods can effectively characterize the vibration of the manipulator's end. Both the vibration amplitude and trend measured by the two methods show strong consistency. Although the vision-based method is more complex in terms of camera setup and calibration, it ultimately avoids the time-consuming process of sensor installation, which is prone to errors. In contrast, IMU methods require precise calibration and frequent maintenance during operation to minimize errors caused by drift effects.

Additionally, vision-based methods can incorporate advanced feature tracking and other techniques to enable real-time monitoring of displacement changes at the end of a large manipulator, offering high accuracy in measuring small vibrations. In contrast, the IMU method struggles to achieve high-precision real-time displacement measurements, especially in the presence of small vibrations in complex environments where significant errors can arise.

In conclusion, the experimental results highlight the significant advantages of the vision-based vibration measurement method in terms of accuracy, stability, and ease of use. This method is particularly suitable for vibration monitoring and maintenance of large hydraulic manipulators, offering robust non-contact measurement capabilities. It shows broad application potential, especially in vibration monitoring scenarios that require high accuracy and stability.

6 Conclusions and outlook

This paper proposes a vision-based, non-contact vibration measurement method, which includes the calibration of the camera's intrinsic and extrinsic parameters, Harris corner detection, and feature point tracking based on optical flow. This study effectively overcomes the limitations of traditional contact-based sensors in complex environments, offering a stable and accurate non-contact vibration measurement solution.

The proposed method is compared with existing contact-based measurement methods (i.e., IMU and angle encoders) in a series of experiments conducted under two different operating conditions. The experimental results demonstrate the vision-based method's superiority in terms of accuracy and stability, accurately extracting displacement data related to manipulator vibrations.

This paper proposes an innovative solution for the vibration monitoring of large hydraulic manipulators. Future research can further optimise the corresponding visual detection algorithm to improve the adaptability and stability of the method in more complex environments. And the accuracy of the proposed method can be further verified by comparing it with other measurement means (laser total station tracking the end of the manipulator).

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. U21A20124).

References

[1] Geimer, M. Mobile Working Machines. SAE International, Warrendale, PA, 2020.

[2] Machado, T., Fassbender, D., Taheri, A., Eriksson, D., Gupta, H., Molaei, A., Forte, P., Rai, P. K., Ghabcheloo, R., Mäkinen, S., et al. Autonomous heavy-duty mobile machinery: A multidisciplinary collaborative challenge. In Proceedings of the 2021 IEEE International Conference on Technology and Entrepreneurship (ICTE), pp. 1–8, Kaunas, Lithuania, August 2021. IEEE.

[3] Try, P., and Gebhard, M. A vibration sensing device using a six-axis IMU and an optimized beam structure for activity monitoring. Sensors, 23(19):8045, 2023.

[4] Patrizi, G., Carratù, M., Ciani, L., Sommella, P., Catelani, M., and Pietrosanto, A. Analysis of inertial measurement units performances under dynamic conditions. IEEE Transactions on Instrumentation and Measurement, 72:3520713, 2023.

[5] Staszewski, W. J., Lee, B. C., Mallet, L., and Scarpa, F. Structural health monitoring using scanning laser vibrometry: I. Lamb wave sensing. Smart Materials and Structures, 13(2):251–260, April 2004.

[6] Baqersad, J., Poozesh, P., Niezrecki, C., and Avitabile, P. Photogrammetry and optical methods in structural dynamics—A review. Mechanical Systems and Signal Processing, 86:17–34, March 2017.

[7] Li, M., Yang, S., He, J., Gu, X., Xu, Y., Gu, F., and Ball, A. D. Full-field extraction of subtle displacement components via phase-projection wavelet denoising for vision-based vibration measurement. Mechanical Systems and Signal Processing, 224:112021, 2025.

[8] Deng, R., Lin, Y., Li, M., Gu, F., and Ball, A. Vision-based vibration mode shape extraction for identifying cracks in a cantilever beam. Journal of Physics: Conference Series, 2184(1):012037, May 2022.

[9] Kalpoe, D., Khoshelham, K., and Gorte, B. V. Vibration measurement of a model wind turbine using high speed photogrammetry. In Proceedings of SPIE, vol. 8085, p. 80850J, June 2011.

[10] Gautama, T., and Van Hulle, M. M. A phase-based approach to the estimation of the optical flow field using spatial filtering. IEEE Transactions on Neural Networks, 13(5):1127–1136, September 2002.

[11] Jepson, A. D., and Fleet, D. J. Phase singularities in scale-space. American Journal of Physics, 67(1):55–60, 1999.

[12] Fleet, D. J., and Jepson, A. D. Stability of phase information. In Proceedings of the IEEE Workshop on Visual Motion, vol. 4, no. 1, p. 1, October 1991.

[13] Lucas, B. D., and Kanade, T. An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence, pp. 674–679, 1981.

[14] Balazadeh Bahar, M. R., and Karimian, G. High performance implementation of the Horn and Schunck optical flow algorithm on FPGA. In Proceedings of the 20th Iranian Conference on Electrical Engineering, pp. 736–741, 2012.

[15] Zhang, Z. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330–1334, November 2000.

[16] Try P , Gebhard M .A Vibration Sensing Device Using a Six-Axis IMU and an Optimized Beam Structure for Activity Monitoring[J].Sensors, 23(19):24, 2023.