Exploratory Study of Drone Data Stabilization with Implications in Vibration-based Structural Health Monitoring

Ngeljaratan, Luna  
Department of Civil and Environmental Engineering, University of Nevada

Mohamed A. Moustafa  
Department of Civil and Environmental Engineering, University of Nevada

Sumarno, Agung  
Research Center for Structural Strength Technology, National Research and Innovation Agency

Agus Mudo Prasetyo  
Research Center for Structural Strength Technology, National Research and Innovation Agency

他

https://doi.org/10.5109/7151727
Exploratory Study of Drone Data Stabilization with Implications in Vibration-based Structural Health Monitoring

Luna Ngeljaratan1*,2, Mohamed A. Moustafa1, Agung Sumarno2, Agus Mudo Prasetyo2, Dany Perwita Sari2, and Maidina2

1Department of Civil and Environmental Engineering, University of Nevada, Reno, USA
2Research Center for Structural Strength Technology, National Research and Innovation Agency, BJ Habibie Science and Technology Center, Gedung 220, Setu, Tangerang Selatan 15314, Indonesia

*Author to whom correspondence should be addressed:
E-mail: luna001@brin.go.id

(Received April 25, 2023; Revised July 23, 2023; accepted July 28, 2023).

Abstract: This exploratory study presents stabilization approaches of drone-based videos with implications in vibration-structural health monitoring applications. Unlike other standalone vision-based systems such as cameras placed on a tripod, drone drifting may occur due to airborne missions especially when operated indoors or in lower altitudes. It affects the captured frames and even the slightest movement of the drone camera will reduce the data accuracy. Post-processing methods using Computer Vision (CV) and signal-processing algorithms are used in this study to explore their effectiveness and accuracy in measuring the dynamic vibration of a structure. The object study is an aluminum bar subjected to two sinusoidal vibrations then recorded using a camera embedded on a quadcopter. After video acquisition, the post-processing is started by image enhancement and scaling procedures, followed by Scale Invariant Transform (SIFT) feature detection, extraction, and matching. The scaling factor is used to convert the image coordinate to an object coordinate system before computing the displacement of the object. Data stabilization techniques are implemented in this study, first is the background subtraction method to eliminate signal drifting and second is cleaning the data from any trends. The accuracy of the proposed framework is tested by comparing the structural dynamic responses from drone measurement to a reference sensor. The Autoregressive (AR) model generated the Power Spectral Density of the signal is also compared to the measurement from a reference camera. The results show the high accuracy of the proposed method which is up to 97.52% on the dynamic response with less effect on the signal PSD. Overall, the exploratory study obtains satisfactory results and provides a new alternative to an intelligent system in the structural health monitoring field.

Keywords: drone-based video, SIFT, stabilization, vibration, structural health monitoring, accuracy.

1. Introduction

The field implementation of vision-based sensor systems for Structural Health Monitoring (SHM) purposes has grown in interest due to their non-contact and remote operations, multiple location measurements, high accuracy, and fast deployment. Therefore, besides focusing on the deployment of vision-based systems in monitoring large-scale infrastructure systems1-18, previous research has also been focusing on their challenges for example estimating their accuracy19,20, recovering, improving, and managing their data21, or testing their capability under challenging environment10. Structural safety assessment also becomes the major contribution of vision-based SHM systems since their data can be further processed to generate valuable information related to structural conditions. Their quick and simple operations provide an alternative option for rapid response and hazard mitigation in which their potentials have been previously studied mostly for seismic monitoring of civil infrastructures.

The market growth of digital cameras supports the aims of vision-based systems for SHM; thus, various types of cameras are used with preferences on a more cost-effective solution. Recent advances in robotics that integrate cameras on their platform add more advantages in remote and safe operation also providing multiple tasks to be executed by the robotic system. Unmanned Aerial Vehicle (UAVs) or commercially known as drone offers advantages in remote and airborne operations since SHM
can be conducted free from ground obstacles such as traffic for bridge dynamic SHM or ground movement for civil infrastructure seismic SHM. Several studies have opened the possibility of using drones for bridge monitoring\(^\text{21}\), pipeline monitoring\(^\text{22}\), or masonry construction monitoring\(^\text{23}\). Challenges were also addressed that mostly focused on drone limitations and automation. Limitation in battery life, for example, causing shorter mission time\(^\text{24}\). Field deployment for leaked inspection is also restricted since it requires additional sensors that add to drone payload capacities\(^\text{25}\). The GPS signal is also an issue when the monitoring is conducted in an indoor environment; thus, it suggests an autonomous drone to avoid a sudden loss of signal during operation\(^\text{26}\).

In vibration-based SHM, the camera embedded in the drone should record steady videos to generate an accurate structural response. Therefore, previous works to stabilize the cameras have been conducted, generally with the support of passive or active components. Passive components such as isolators have been developed to dissipate energy\(^\text{27,28}\) and an advanced isolation system using viscoelastic materials has been proposed to stabilize drone camera\(^\text{29}\). Active components are using additional sensors to stabilize the drone camera\(^\text{30}\) or a piezoelectric actuator\(^\text{31,32}\) to dissipate drone vibration from the camera. Similar goals of drone stabilization for vibration-based SHM are also shared with this work with different approaches using image processing and mathematical model. This effort belongs to the passive approach that relies on the post-processing of drone videos. Without using any mechanical equipment, no additional payload is required; therefore, the proposed technique compensates for the issues of drone battery life and mission time.

The general concept of the work is drone for civil infrastructure seismic safety assessment as shown in Figure 1. The drone system is developed purposely for seismic monitoring of large-scale civil infrastructure with high-risk categories such as bridges or buildings in risk category IV as depicted by ASCE 7-16. The drone system records the global response of the structures then the changes in structural dynamic characteristics are analyzed and reported. Structural identification investigates the seismic performance of the structures resulting in a structural safety assessment that confirms the final state of the structure after the major seismic events.

![Fig. 1: Drone for seismic structural health monitoring—general concept.](image)

The drone concept for seismic SHM as given in Figure 1 is elaborated into several tasks in which the preliminary task is covered in this study which is to explore the capability of a drone for vibration-based SHM highlighting the drone's nature movement and drifting. The objective of this study is to explore the data stabilization method and algorithms to eliminate drone movement and to implement the results into vibration-based SHM. The methods briefly explain the drone axes and rotation that naturally result in drone movement and drift, the framework for drone vibration SHM considering data correction, a brief review of video and image processing of drone data using several computer vision algorithms, and structural identification method to estimate the dynamic characteristic of the structure. The framework is tested first on an aluminum bar placed on a simulator that generates two sinusoidal signals before implementation on a larger scale structure. A quadcopter is operated and recorded the response of the bar and the data are further analyzed using the methods described previously. The accuracy is estimated by comparing the result of the first test with a reference camera and the conclusion is summarized at the end of this paper.

### 2. Methods

#### 2.1 Drone axes and rotation

Prior studies have explored drone parameters\(^\text{33-38}\) such as their integrity, identification, and performance. Similar to a commercial airplane, the drone also maneuvers in three directions as shown in Figure 2. These three axes, namely the lateral or transverse axis, longitudinal axis, and vertical axis, are each perpendicular to each other and intersect at the drone's center of gravity. During the flight, a drone is capable of using these axes and in turn controls its direction by powering the rotors and changing the drone speed. Motion around the lateral axis, longitudinal axis, and vertical axis are referred to as pitch, roll, and yaw, respectively. Pitch motion originates at the middle of the drone's nose and drops the tail. The roll motion also begins at the middle of the drone's center of gravity and it is directed forward, parallel to the fuselage reference line. A positive roll motion lifts the drone's nose and drops the tail. The yaw motion turns the nose of the drone to the right.

When a drone is on a mission, internal and external factors may contribute to its drift and movement. Failed calibration, issues on the motor, damaged propellers, or unequal weight distribution as the drone carries other sensors are several internal factors that cause the drifting. The environmental condition also affects drone stabilization, for example, extreme wind is an issue for small drone and electromagnetic interference disturb the navigation system that disrupts the communication system between drone and controller, causing the drone to operate
intensely that may cause it to crash into any obstacles. For small drone movement and drifting, the resulting time signal or series such as structural displacement may show a specific trend that requires processing to remove and stabilize the signal. These are background of this study that such challenges can be minimized by providing an effective solution to stabilize drone data without adding more instrumentations on the drone platform that only adds drone payloads.

2.2 Framework

The framework of the drone for vibration SHM is shown in Figure 1 which provides several steps from video to data processing with the final goal of structural identification. Each step is elaborated and this framework is tested and validated using small vibration tests using a simulator that is further described in the next section.

2.2.1 Video and image processing

The drone mission in this study records a video for each test with a low sample rate of 30 Hz (which means the drone records 30 frames per second). The selection of the frame rates is based on the previous works\(^{3,9-11}\) that successfully capture the natural frequency of the tested large-scale structures using only a 30 Hz sampling rate. The video is then transformed into images and these images are further processed using image processing algorithms to obtain clear images without motion blur and low brightness. Most computer vision algorithms work only when the object is illuminated adequately since their basic is the pixel intensity map that relies dependently on light distribution. The classical histogram equalization (HE) that is used in other disciplines\(^{39-41}\) also is adopted in this study to process low-contrast images recorded by drone as shown in equation (1) in which \(I\) and \(\Theta\) are the input and output images. The transform function \(\mathcal{F}\) in equation is based on the cumulative density function (CDF) that maps the input image \(I(x, y)\) into the entire dynamic range \((I_L, I_H)\). To accelerate the image processing process, an image compression algorithm is also adopted in this study in which the compression ratio of two is used. A brief overview of the image compression algorithm is expressed in equation (2) with an explanation as follows. Considering a pixel \((x, y)\) presenting a horizontal and vertical direction, respectively, is located on an image \(I\) with a size of \((X \times Y)\). As resampling factors of \(r_x, r_y\) \((r_x = r_y)\) are applied on image \(I\), image \(I\) is repositioned by rotating it by angle \(\theta\). Consequently, the new coordinate of pixel \((x, y)\) is now \(j(x', y')\) that is computed by adding a translation vector \(t_x, t_y\) in each direction of the image.

\[
O(x, y) = \{\mathcal{F}(I(x, y)) \forall (x, y) \in I\}
\]

\[
\begin{bmatrix}
x' \\
y'
\end{bmatrix} = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix} \begin{bmatrix}
x \\
y
\end{bmatrix} + \begin{bmatrix}
t_x \\
t_y
\end{bmatrix}
\]

2.2.2 Feature detection, extraction, and matching

Several object recognition algorithms have been proposed to provide fast and accurate feature detection, extraction, and matching. This study adopts the Scale Invariant Transform (SIFT) algorithm that has been extensively used for image classification tasks. The SIFT procedures include the Difference of Gaussians (DoG) Space Generation, Key points Detection, and Feature Description. The DoG process is generally divided into smoothing the initial image using Gaussian kernels and down sampling the processed image using Gaussian Scale Space. In the detection of the Key point, the point that is defined as a key point needs to be compared with its neighborhood (usually eight neighborhoods) at their similar layers and nine neighbors for the upper and lower layers. Then, SIFT feature is described as the gradient magnitude and gradient direction around the key points. More details about SIFT features can be found in\(^{42}\). Illustration of how SIFT feature work between two images is shown in Figure 3. After the detection and extraction of SIFT features in both images, they are then matched and the unmatched features should be eliminated using several threshold values. Figure 3 shows an excellent matching feature between the images, there is no miss-match found between image 1 and image 2.

2.2.3 Data correction

After SIFT feature detection, extraction, and matching, the location of each feature is tracked among the consecutive images resulting in pixel coordinates in horizontal and vertical directions since the proposed framework is specifically tested for two-dimensional directions. A scale factor is used to transform into displacement time-series data in the unit of mm. When the
drone experiences movement or drift, its motion can be modeled as shown in Figure 4 which demonstrates vertical and horizontal motions versus the movement of the drone camera. The model is based on the SIFT feature traced from the artificial targets attached to the drone body and camera. It will impact the measurement data that will be explained further in this next section.

To eliminate drone movement and drift and to stabilize the measurement data, two simple techniques are implemented in this work. The first technique is to extract the steady feature such as a background feature, i.e. SIFT feature detected from the object background. When the drone is assumed to experience drift, the object data recorded by the drone are the combination between the real vibration plus the drone movement. However, as for steady background feature data recorded by drone, the data only show the drone movement itself as the feature does not move during the test. Therefore, it is easier to obtain the real object data, RO', by extracting the raw object data, RO, and the background feature data, BG. After background subtraction, the data are examined if a specific trend exists such as a linear or quadratic trend. Data are cleaned by removing these trends as the second technique implemented in this study. Finally, the accuracy of the proposed method is evaluated by measuring the Root Mean Square Error (RMSE) between the drone's final data and a reference sensor.

2.2.4 Structural identification

After data correction and accuracy measurement of drone final data to a reference sensor, the final step is to characterize the dynamic properties of the tested object using a structural identification method. One of the most common methods besides decomposition techniques \(^4\) or neural networks to extract the fundamental frequency of a signal using spectrum estimation is a Fast Fourier Transform (FFT) \(^4\). However, extracting fundamental frequency using the FFT method required long-duration data records to generate a correct frequency resolution. Therefore, this work adopts the Auto-Regressive (AR) model for spectral estimation that offers an alternative to the FFT method. \(^4\) It is a parametric model that provides a greater spectral resolution as compared to its other parametric model counterparts and also yields reasonable spectral estimates for short data records. The AR coefficients in this study are computed using the Covariance method. \(^4\) The AR covariance process of order \(p\) is expressed in equation (3) in which a white noise with variance \(e\) is filtered and resulted in a stationary random process \(y(n)\). The power spectral density of \(y(n)\) is defined as \(P_y(e^{j\omega})\) that strongly depends on the selection of order \(p\).

\[
P_y(e^{j\omega}) = \frac{e}{1 + \sum_{j=1}^{p} a(j)e^{-j\omega}}
\]

3. Experimental validation

3.1 Testing setup

The proposed framework was experimentally evaluated through two-sine wave tests. A small simulator was used to generate sinusoidal loading for a rigid test model at the Earthquake Engineering Laboratory at the University of Nevada, Reno. The model was an aluminum block with the dimensions \(100 \text{ mm} \times 10 \text{ mm} \times 900 \text{ mm}\) as shown in Figure 4b. The monitoring is conducted by a drone-quadcopter-type shown in Figure 4a recording the movement of the three artificial targets, numbered T1, T2, and T3, placed on the specimen shown in Figure 4b. The background target (BG) shown in Figure 4b is used to eliminate the drone drifting through a background subtraction technique explained in the previous section. Those targets were printed on adhesive stickers and permanently attached to the specimen. The radius of the white circular targets was \(6.5 \text{ mm}\).

Before the test, the gimbal and cameras were first calibrated and the camera setting was adjusted to record the best video quality. For both tests, monitoring was set at \(30 \text{ frame-per-second (fps)}\), i.e. \(30 \text{ Hz}\) sampling rate, and utilized the full resolution ROI of the drone cameras, i.e. \(3840 \times 2160 \text{ pixels}\). A partial specification of the drone is given in Table 1. After calibration, the drone was placed on the helipad shown in Figure 4a preparing to take off. Using the controller, the drone was navigated to position in front of the specimen with the sample view of the drone recording videos during the tests shown in Figure 4c. Then, the test was started by running the simulator driven by the controller shown in Figure 4b that applied a prescribed sinusoidal pattern to the specimen. After the first test, the drone was navigated back to the helipad and prepared for the second test. This step concluded the drone data collection and was continued by the data processing following the framework in Figure 2. After video-to-image conversion, the images were compressed with the compression ratio of two giving the resolution of the processed image as \(1920 \times 1080 \text{ pixels}\) to accelerate the process. For validation purposes, a smartphone camera placed on a fixed tripod was used to monitor the first, and the results between the two systems were compared. The displacement measurements from both tests are reported and the frequency content is computed as discussed in detail in the next section.
Table 1. Drone specifications

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aperture</td>
<td>f/2.8 - f/11</td>
</tr>
<tr>
<td>Battery Life</td>
<td>30 Minutes, 5870 mAh</td>
</tr>
<tr>
<td>Flight Range</td>
<td>7 km</td>
</tr>
<tr>
<td>Gimbal stabilization</td>
<td>3 axes</td>
</tr>
<tr>
<td>ISO</td>
<td>100 - 6400 (Video)</td>
</tr>
<tr>
<td></td>
<td>100-12800 (Photo)</td>
</tr>
<tr>
<td>Shutter Speed</td>
<td>8 - 1/2000 s (Mechanical)</td>
</tr>
<tr>
<td>Satellite Positioning</td>
<td>GPS</td>
</tr>
<tr>
<td>Systems</td>
<td></td>
</tr>
<tr>
<td>Video Transmission</td>
<td>720p</td>
</tr>
<tr>
<td>Weight</td>
<td>1288g</td>
</tr>
<tr>
<td>Size</td>
<td>350mm</td>
</tr>
</tbody>
</table>

Table 2. Displacement measurement error from corrected data to reference sensor.

<table>
<thead>
<tr>
<th>Error</th>
<th>Test #1</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_{\text{max}}$</td>
<td>mm</td>
<td>6.52</td>
<td>6.54</td>
<td>6.47</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>2.45</td>
<td>2.48</td>
<td>2.43</td>
</tr>
<tr>
<td>$e_{\text{avg}}$</td>
<td>mm</td>
<td>3.07</td>
<td>3.04</td>
<td>3.06</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>1.14</td>
<td>1.13</td>
<td>1.15</td>
</tr>
<tr>
<td>RMSE</td>
<td>mm</td>
<td>3.31</td>
<td>3.35</td>
<td>3.24</td>
</tr>
</tbody>
</table>

3.2 Displacement response

This section presents the comparison and error analysis of drone displacement measurements with validation from Test 1 which is compared to a reference sensor, i.e. a smartphone. The direction of the applied load to the drone camera orientation was in the horizontal, i.e. in-plane direction, or is defined in this work as $x$-direction. The displacement of the specimen is calculated by subtracting the displacement of the attached targets, i.e. T1, T2, and T3 to BG, all measured from drone videos.

The accuracy of displacement measurement results and errors is estimated for Test 1 from three artificial targets, i.e. T1, T2, and T3. Figures 5 shows the displacement comparison measured from Test 1. The first observation is that a linear trend is present in displacement data. To further analyze these observations, the mean displacement values $\mu_{T1}$, $\mu_{T2}$, and $\mu_{T3}$, are removed from the test data. The corrected data are then plotted together with the displacement measured from the reference sensor.

The comparison of the maximum error, $e_{\text{max}}$, and average error, $e_{\text{avg}}$, for each of the three targets are summarized and presented in Table 2 for Test 1. Since this study emphasizes quantifying the accuracy of drone measurement after the stabilization or correction process, the errors presented here are both in engineering units, i.e. mm, and normalized percentage. The table shows the maximum error between the corrected drone displacement is within the range of 6.4-6.5 mm that is associated with the percentage of 2.4-2.5%. The peak is estimated from T2 data as 6.54 mm or 2.48%. This result implies the drone measurement accuracy of 97.52%, with respect to the measurement from reference sensor. Another accuracy measurement is obtained from the RMSE value, that relatively shows small difference among the three targets with the largest value computed as 3.35 mm.

Another variation of drone data is presented in Figure 7 from Test 2. If linear trend is observed from Test 1 data, a quadratic trend is shown based on the plot from Test 2. The original data measured from the three targets as shown in Figure 7, indicate a trend in a certain polynomial degree $n$. Using a quadratic $n = 2$, the trend is removed and the corrected data are plotted together in Figure 7.

3.3 Frequency response

Since no reference sensor is deployed for the second test, further verification is conducted in the frequency domain assuming that, since the applied load and response both show sinusoidal waveforms, then the computed frequency contents between both tests should be closer or nearly identical. In this section, the expected frequency content of the specimen is presented using the AR-
covariance algorithm for three targets attached on the specimen as previously shown in Figure 5. The frequency response of the specimen presented here is in a healthy state (no damage) due to the applied sinusoidal waveforms. These waveforms are important in studying seismic response since most structures that produce resonance will generate a sinusoidal motion. Also, these input waves have low amplitudes so they cause no harm to the specimen; thus, it is ideal for identification in a damaged state. The input waves applied to the specimen were a 30-sec duration. For this analysis, the AR-covariance algorithm was used to extract the specimen’s natural frequency. A spectral estimation using AR method has advantages over Fourier transform as instead of estimating the power over a fixed-range, the method can work over a continuous range of frequencies. Also, less variance is observed on the power spectral estimation when AR method is selected.

For the verification purpose, the acceleration data from Test 1 was processed and resulted in a PSD spectrum as shown in Figure 8, similar to the verification conducted for the displacement data. As for the identification from the second test, the accelerations from similar targets were analyzed using a similar AR-covariance algorithm. Results of the natural frequency of the specimen are represented below and compared with each other.

Figure 8 shows the longitudinal response of the specimen recorded at three targets using the acceleration record from Test 1. Similar to displacement response, the spectrum is plotted for the original, corrected, and reference data. The total length of the record is 29 sec with 872 data points and a sampling rate of 30 Hz. It shows the results of the spectral estimation with the obtained eigenfrequencies indicated at the peak amplitude.

The AR spectra of three different targets analyzed from Test 1 is shown in Figure 8. Though in displacement-time domain a slightly small difference is observed among targets, in the frequency domain, similar peaks at 5.27 Hz, 5.20 Hz, and 5.23 Hz for their original, corrected, and reference data are computed. Meanwhile, Test 2 produces a slightly higher amplitude of PSD as shown in Figure 9. The peaks are shown at 5.27 Hz, 5.51 Hz, and 5.37 Hz for targets 1, 2, and 3, respectively. These results differ by approximately 4.25% for the maximum frequency and by 1.35% when measured by the minimum values from Test 1 and Test 2.

4. Conclusion

This paper presents an exploratory study of drone data stabilization with the implementation of vibration structural health monitoring. The proposed framework is tested on an aluminum bar vibrated using sinusoidal waveforms. SIFT feature detection, extraction, and matching are first applied to identify three artificial targets attached to the specimen. Two drone data stabilization are proposed, first a stationary background extraction is suggested to eliminate drone drift. Second, a simple mathematical model is suggested to remove trends present in the original data. Verification is conducted to analyze the error between the corrected data to a reference sensor. The following are concluding remarks from this study:

- A comparable result is obtained from drone-corrected data to reference sensors that provide confidence to further use the proposed framework for drone data stabilization specifically in the structural health monitoring field. The maximum error in the displacement domain is estimated as 2.48% associated with the displacement accuracy of 97.52%.

- Less variation in the frequency domain is observed among targets. The AR-covariance method results in the estimated natural frequency within the range of 5.2% from Test 1 and it returns very comparable values between the corrected and reference data. Overall, the proposed framework for drone data stabilization can be successfully used for vibration structural health monitoring as demonstrated from the two tests conducted in this study. Future work may consider the use of commercial lower-end, DSLR, or surveillance cameras to further validate the techniques for vibration-based structural health monitoring using drones, especially when drone stabilization becomes a concern in field deployment.

Acknowledgments

This work was funded by Accelerated Bridge Construction University Transportation Center (ABC-UTC), USA. We also thank The Indonesia Endowment Funds for Education (LPDP) and the National Research and Innovation Agency (BRIN) through Riset and Inovasi untuk Indonesia Maju (RIIM) program (Grant Number
References


19) Ngeljaratan L, Moustafa MA. Uncertainty and Accuracy of Vision-based Tracking Concerning Stereophotogrammetry and Noise-floor Tests 2022;


25) Bretschneider TR, Shetti K. UAV-based gas pipeline leak detection. 2015;


32) Oh J-S, Han Y-M, Choi S-B. Vibration control of a camera


