Special Issue Article

Accurate vision-based displacement and vibration analysis of bridge structures by means of an image-assisted total station

Mohammad Omidalizarandi(iD), Boris Kargoll(iD),
Jens-André Paffenholz(iD) and Ingo Neumann

Abstract

Today, short- and long-term structural health monitoring of bridge infrastructures and their safe, reliable and cost-effective maintenance have received considerable attention. For this purpose, image-assisted total station (here, Leica Nova MS50 MultiStation) as a modern geodetic measurement system can be utilized for accurate displacement and vibration analysis. The Leica MS50 measurements comprise horizontal angles, vertical angles and distance measurements in addition to the captured images or video streams with practical sampling frequency of 10 Hz using an embedded on-axis telescope camera. Experiments were performed for two case studies under (1) a controlled laboratory environment and (2) a real-world situation observing a footbridge structure using a telescope camera of the Leica MS50. Furthermore, two highly accurate reference measurement systems, namely, a laser tracker Leica AT960-LR and a portable shaker vibration calibrator 9210D in addition to the known natural frequencies of the footbridge structure calculated from the finite element model analysis are used for validation. The feasibility of an optimal passive target pattern and its accurate as well as reliable detection at different epochs of time were investigated as a preliminary step. Subsequently, the vertical angular conversion factor of the telescope camera of the Leica MS50 was calibrated, which allows for an accurate conversion of the derived displacements from the pixel unit to the metric unit. A linear regression model in terms of a sum of sinusoids and an autoregressive model of the coloured measurement noise were employed and solved by means of a generalized expectation maximization algorithm to estimate amplitudes and frequencies with high accuracy. The results show the feasibility of the Leica MS50 for the accurate displacement and vibration analysis of the bridge structure for frequencies less than 5 Hz.

Keywords

Image-assisted total stations, calibration, autoregressive model, generalized expectation maximization, displacement and vibration analysis, footbridge monitoring

Introduction

Today, short- and long-term structural health monitoring (SHM) of the bridge infrastructures and their safe, reliable and cost-effective maintenance has received considerable attention. For this purpose, various measurement systems with different levels of accuracies and

Geodetic Institute, Leibniz Universität Hannover, Hannover, Germany

Corresponding author:
Mohammad Omidalizarandi, Geodetic Institute, Leibniz Universität Hannover, Nienburger Straße 1, 30167 Hannover, Germany.
Email: zarandi@gih.uni-hannover.de

prices are being widely used depending on the demand. SHM is commonly being conducted based on visual observation, the properties of the material of the structures and the interpretation of the structural characteristics by inspecting the changes in the global behaviour of the structure (e.g. natural frequencies, mode shapes and modal damping).[1] Therefore, SHM is interpreted as a process to detect structure damages or identify their characteristics by the discrete or continuous measurements over time. Furthermore, the dynamic characteristics of a structure, such as frequencies, can change due to the temperature variations or the damages occurring in the structure.[2] From a surveying engineer's point of view, it is crucial to detect any deterioration of the structures (even small cracks) by frequent measurements. Typically, the geodetic measurement systems, such as total station, robotic total station (RTS), terrestrial laser scanner (TLS), laser tracker (LT), global positioning system (GPS) or other sensors such as digital camera or accelerometer, can be used for displacement and vibration monitoring. The Nyquist theorem must be fulfilled to identify the frequencies of the oscillating structure correctly. Consequently, the proper measurement systems must be used according to the sampling frequency required and the maximum amplitude derived from the oscillation of the structure.[3] Bridges (including footbridges or road bridges) generally oscillate in a range of 1.2–10 Hz (or more).[4] Previous researchers used different geodetic sensors for vibration monitoring of the bridge structures. Psimoulis and Stiros,[5] for example, used RTS for vibration monitoring of a cable bridge, pedestrian suspension bridge and steel railway bridge for non-constant sampling rate measurements of the RTS in a range of 5–7 Hz and performing spectral analysis based on the Norm-Period code.[6] Roberts et al.[7] presented the hybrid configuration of GPS with a sampling frequency of 10 Hz and triaxial accelerometer with sampling frequency of 200 Hz for a bridge deflection monitoring. On one hand, the accelerometer measurements of this hybrid measurement system were beneficial to eliminate the disadvantages of GPS measurements regarding multipath, cycle slips errors and the need for good satellite coverage. On the other hand, GPS measurements were utilized to suppress accumulation drift of the acceleration data over time through velocity and coordinate updates. Neitzel et al.[8] used the sensor network of the low-cost accelerometers with a sampling frequency up to 600 Hz, the TLS (Zoller + Fröhlich Imager 5003) in a single-point measurement mode with a sampling frequency of 7812 Hz and a terrestrial interferometric synthetic aperture radar (t-InSAR) with a sampling frequency of 200 Hz for vibration analysis of the bridge structure. They defined a functional model based on a damped harmonic oscillation and solved it in the sense of the least square adjustment. The reason for such a high sampling frequency of the TLS was to detect displacements smaller than 1 mm by averaging the measurements over 100 measurements and to reach a practical sampling frequency of 78.12 Hz. An overview of the TLS-related structural monitoring was given in Vosselman and Maas.[9]

Structural monitoring by means of the vision-based measurement technologies is becoming increasingly popular in the context of civil engineering structures such as buildings, bridges and dams. In particular, an image-assisted total station (IATS), which can be a total station with an integrated external camera, for example, a high-resolution digital camera mounted on top of the scanning system by means of a clamping system, cf. Omidalizarandi et al.,[10,11] or an internally embedded camera, was employed by Reiterer et al.,[12] Bürki et al.,[13] Wagner et al.,[14,15] Ehrhart and Lienhart,[2,16] Guillaume et al.,[17] Wagner[18,19] and Lienhart et al.[3] Most modern IATS measurement systems have motorized axes of rotation, which allow for an automatic rotation of the telescope to the points previously measured at different epochs of time. The IATS measurements comprise horizontal directions, vertical angles and distance measurements (in a polar coordinate system) in addition to the captured object images or video streams using embedded or externally attached cameras. The internally embedded on-axis telescope camera in addition to the motorized axes of rotations is particularly well-suited to an accurate, automatic and autonomous measurement of structures in static and dynamic monitoring. Subsequently, it enables us to measure both active targets (i.e. retro-reflective prism targets) and passive targets (i.e. signalized or non-signalized targets). Therefore, IATS with an on-axis telescope camera is advantageous over other vision-based measurement systems since the displacements in the image space can be converted directly to the metric unit by means of total station capabilities. In addition, its stability over time can be controlled by measuring its telescope angles and tilt reading[3] using GeoCOM interface.[20]

Ehrhart and Lienhart[16] used the telescope camera of an IATS for the displacement and vibration monitoring of a footbridge structure based on the captured video frames of the circular target marking rigidly attached to the structure. Afterwards, least-squares ellipse fitting based on the Gauss–Helmert model (GHM) was applied to extract the target centres. Ehrhart and Lienhart[2] and Lienhart et al.[3] employed an IATS (Leica MS50 with a sampling frequency of 10 Hz), an RTS (Leica TS15 with a sampling frequency of 20 Hz) and an accelerometer (HBM B12/200 with a sampling frequency of 200 Hz) for vibration analysis of a footbridge structure based on measurements of the circular target markings (i.e. signalized targets) and structural features (like bolts, that is, non-signalized targets) of the bridge.

The choice of a feasible optical target pattern and its accurate, automatic recognition at different epochs of time is the preliminary step in image-based structural monitoring using passive targets. Different target patterns with different detection techniques were previously proposed by several researchers, for instance, least-squares template matching by Gruen,[21] Akca,[22] Gruen and Akca[23] and Bürki et al.;[13] coded target detection by Zhou et al.;[24] ellipse detection by Ehrhart and Lienhart[16] and Guillaume et al.;[17] circle matching by Bürki et al.;[13] cross-line detection by Reiterer and Wagner[25] and centre-of-mass detection by Bürki et al.[13] Omidalizarandi et al.[26] presented an optimal circular target with the line pattern consisting of four intersected lines and proposed a target centroid detection approach, which was shown to be robust, reliable and accurate regarding the lighting condition, dusty environment and skewed angle targets.

After extracting the target centroid, the next step is to accomplish camera calibration to convert the pixel (px) coordinates to more meaningful metric quantities, such as theodolite angle readings (i.e. horizontal and vertical directions[27–29]). Based on the pixel differences between the initial pointing to the corresponding target of interest and the precisely calculated direction to the detected target centroid, an accurate remeasurement of the centroid of the target is also possible (i.e. taking advantage of the motorized axes of rotation of the IATS). However, the instrument's axes errors, vertical-index error and collimation error can also be considered to perform the conversion from the pixel to the metric coordinates more precisely. In order to capture sharp images with the telescope camera of the IATS, the corresponding targets should be focused by turning on the autofocus capability of the IATS. This, however, leads to changes in the internal camera calibration

parameters. Zhou et al.[24] proposed IATS telescope camera calibration based on measurements of the coded targets and their angular reading from the total station at a certain focus positions. Subsequently, new sets of calibration parameters were calculated by means of cubic polynomial interpolation at certain focus positions. In the context of displacement monitoring and for the purpose of converting object movements within the image space from the pixel unit to a proper angular quantity, Ehrhart and Lienhart[16] merely calibrated the vertical angular conversion factor in the temperature-controlled laboratory with a fixed and stable set-up of the total station and the target. The video frames of the circular target marking were captured with a sampling frequency of 10 Hz at a fixed position for different horizontal and vertical rotations of the telescope camera of IATS. In addition, telescope angles were measured with a sampling frequency of 20 Hz to improve the measured reference angles by averaging.

The focus of this research is to perform accurate displacement and vibration analysis for two case studies under (1) a controlled excitation in a laboratory environment and (2) an uncontrolled excitation in a real-world situation observing a footbridge structure using the telescope camera of the Leica MS50 (Figure 1, left) with a sampling frequency of 10 Hz. Furthermore, two highly accurate reference measurement systems, namely, a laser tracker Leica AT960-LR (with a sampling frequency of 200 Hz; Figure 1, middle) and a portable shaker vibration calibrator (PSVC) 9210D (with a sampling frequency of 200 Hz; Figure 1, right) in addition to the known natural frequencies of the footbridge structure calculated from the finite element model (FEM) analysis are used for validation. To perform accurate displacement and vibration analysis, first, the feasibility of the optimal passive target pattern and its

Figure 1. Leica Nova MS50 MultiStation (left), Laser Absolute Tracker AT960-LR (middle) and Portable Shaker Vibration Calibrator 9210D (right).

accurate and reliable detection at different epochs of time are investigated. Subsequently, the vertical angular conversion factor of the telescope camera of the Leica MS50 is calibrated, which allows for an accurate conversion of the derived displacements from the pixel unit to the metric unit. A linear regression model in terms of a sum of sinusoids and an autoregressive (AR) model of the coloured measurement noise is employed to estimate amplitudes and frequencies with high accuracy. The white noise components of the AR process are assumed to independently follow a scaled (Student's) *t*-distribution to accommodate for outliers. The adjustment of this combined observation model is carried out by means of the generalized expectation maximization (GEM) algorithm described in Alkhatib et al.[30] In the first application and under controlled excitation, we compare the oscillation frequency and the amplitude derived from the PSVC time series with the results obtained from the Leica MS50 video frames and the LT. In the second application and under uncontrolled excitation, we compare the oscillation frequency and the amplitude derived from the LT time series with the Leica MS50 video frames of the footbridge structure.

Sensor specifications and measurement systems

We used an IATS (here, Leica Nova MS50 MultiStation; Figure 1, left) for displacement and vibration analysis in our experiments. The Leica MS50 includes the following:

1. Precise three-dimensional (3D) laser scanning;
2. A precise total station with an
   - Angular accuracy of 1″ (according to ISO 17123-3);
   - Optical-distance measurement accuracy of 1 mm + 1.5 ppm (according to ISO 17123-4) for prism targets from 1.5 to 10,000 m; as well as
   - Optical-distance measurement accuracy of 2 mm + 2 ppm for non-prism targets (i.e. here passive targets) from 1.5 to 2000 m with a measurement time of 1.5 s;
3. An overview camera with diagonal field of view (FOV) of 19.4°;
4. A telescope camera with diagonal FOV of 1.5°;
5. GNSS connectivity.

Both the overview and the telescope cameras include 5 MP complementary metal-oxide semiconductor (CMOS) sensors in which the telescope camera is an on-axis camera located on the optical path of the Leica MS50 with 30× optical magnification of the overview camera.[31] In this work, we benefit from the total

station's capabilities in terms of precise distance measurements of the passive targets, in addition to the digital imaging by means of the telescope camera. The angular resolution (α) of the telescope camera is approximately 1.7″/px, which is basically calculated by dividing the diagonal FOV by the diagonal length of the captured image in the pixel unit. To accomplish displacement monitoring more accurately, the angular resolution should be calculated from the calibration procedure in a controlled laboratory environment. According to Leica Geosystems,[31] the horizontal and vertical FOVs of the telescope camera of the Leica MS50 are 1.3° and 1.0°, respectively. As we calculate the vertical displacement in our experiments, we merely benefit from its vertical angular conversion factor for displacement and vibration monitoring based on the live video stream functionality of the Leica MS50. The nominal sampling rate of the live video stream of the Leica MS50 is 20 Hz.[31] However, in practice, we could capture the video stream with a sampling frequency of 10 Hz using OpenCV library. In order to get full access to the individual functionality of the Leica MS50 and for ease of use, we made use of the GeoCOM interface,[20] which is written in the script language Python 3.4. In addition, the target centroid detection algorithm proposed by Omidalizarandi et al.[26] is utilized to extract the target centroid with high accuracy. The resolutions of the captured images vary from 320 × 240 px to 2560 × 1920 px. However, for live video stream, it is only possible to capture video streams with a resolution of 320 × 240 px. Subsequently, the target centroid detection approach should be robust, reliable and accurate and should work well even in the case of a low-resolution image. The autofocus is set to 'on', and the white balance is set to 'automatic' to ensure the capture of sharp images. We performed vibration analysis with an image resolution of 320 × 240 px and 1× zoom (Figure 2, left); however, we achieved meaningless results, since such small displacements were not detectable at all. Therefore, as Ehrhart and Lienhart[16] proposed, we reduced the FOV by 8× zoom using the *camera zoom factor* functionality of the GeoCOM interface (Figure 2, right), which gave us reasonable results.
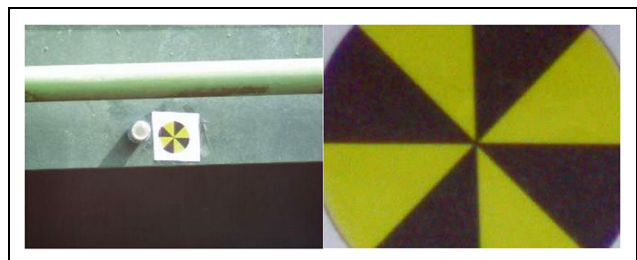
Figure 2. The Leica MS50 telescope camera images: image resolution of 320 × 240 px and 1× zoom (left) and image resolution of 320 × 240 px and 8× zoom (right).

It should be noticed that reflectorless distance measurements to the targets allow us to set the telescope's focus motor position precisely.[16] The entire procedure is controlled via our self-developed graphical user interface (GUI), which allows an efficient and effective data acquisition and analysis.

Two highly accurate reference sensor systems are utilized to perform validation. On one hand, a LT (Figure 1, middle) with a maximum permissible error of $15 \mu m + 6 \mu m/m$ of a 3D point data and measuring rate output of 3000 points per second (3000 Hz)[32] was employed. It allows for sub-millimetre range accuracy of the target points, which can be considered as a reference coordinate frame. On the other hand, a PSVC (Figure 1, right) was employed to perform controlled excitation. It consists of a highly accurate reference accelerometer (in our case, a precise PCB ICP quartz reference accelerometer) and two sensitive dials, which allow us to adjust the frequency and amplitude. For a frequency in the ranges of $0.7 Hz$–$2 kHz$ and $2 Hz$–$2 kHz$, the acceleration can be read out with accuracies of $\pm 3\%$ and $\pm 10\%$, respectively.[33]

Passive target centroid detection

The displacement time series for the captured video frames from the telescope camera of the Leica MS50 can be generated based on the continuous extractions of the point features (i.e. being signalized or non-signalized) at different epochs of time. Omidalizarandi et al.[26] proposed an optimal passive target (Figure 3) and its centroid detection approach to tackle this problem. The proposed target constitutes a circular border with line pattern including four intersected lines. It is low-cost and easy to mount. In addition, its target centroid detection approach is accurate, automatic and fast as well as robust and reliable regarding skew angle

Figure 3. Designed passive target.

targets and poor environmental conditions, such as low lighting (i.e. which may be very bright, semi-dark and dark) and dusty situations. However, the detection approach failed in totally dark lighting conditions.

The procedure starts by manual initial pointing of the targets of interest, which is only carried out at the beginning of the measurements. The telescope is then rotated automatically by means of the motorized axes of rotations of the Leica MS50 to the stored positions of the corresponding targets, and images are captured by means of the telescope camera. Subsequently, target centroid detection (see Algorithm 1) is applied, and the telescope is rotated automatically to the detected target centroid to capture images or video frames.

Cropping of the captured images is carried out, taking into account the target object size (in our case 0.06 m), slope distance (here, the maximum slope distance is up to 30 m) and horizontal as well as vertical FOVs of the telescope camera to extract relevant line features of the aforementioned target pattern and to speed up the procedure. Horizontal and vertical FOVs are calculated according to

$$FOV = 2 \cdot \tan^{-1}\left(\frac{D}{2 \cdot S}\right) (rad) \qquad (1)$$

where D is the target size in the object space and S the slope distance in metres. The calculated FOVs are multiplied by the width and height of the captured image to obtain the width and height of the cropped image. For practical reasons, the width and height of the cropped image are considered three times larger than values calculated to cover the target and its surroundings. This has been found to be beneficial for extracting the target centroid in a significant displacement of the target. We discarded the localising of the target pattern in the captured images by assuming a good initial target pointing at the beginning of the measurement. For further information concerning the localisation of target, please refer to Omidalizarandi et al.[26] The median blur and bilateral filtering are applied to reduce the noise and preserve the sharp edges of the images, respectively.

The line segment detector (LSD) algorithm[34] is applied with stable threshold parameters to extract the line features:

- The sigma value of the Gaussian filter is set to 0.75;
- The bound of quantization error on the gradient norm is set to 2.0;
- The gradient angle tolerance is set to 22.5°;
- The minimal density of region points in the rectangle is set to 0.7;
- The number of bins is set to 1024;
- The gradient modulus in the highest bin is set to 255.

Algorithm 1: Target centroid detection using telescope camera of the Leica MS50

Initialisation:
  initial pointing to the target
Input:
  D = target size in object space (m)
  S = slope distance (m)
  captured image
Output:
  Detected target centroid (px)
Procedure:
  w = image width
  h = image height
  $FOV = 2 \cdot \tan^{-1}\left(\frac{D}{2 \cdot S}\right)$
  $d_w = 3 \cdot w \cdot FOV$
  $d_h = 3 \cdot h \cdot FOV$
  crop captured image using d_w and d_h
  median blur and bilateral filtering
  edge detection using LSD
  compute line azimuths
  calculate histogram of line azimuths
  sort azimuth bin centres in descending order
  While (azimuth bin centres array is empty?)
    select azimuth bin centre from beginning
    select all parallel lines
    convert line vectors to raster
    RANSAC line fitting
    Huber line fitting
    discard selected azimuth bin centre
  End
  compute intersection points of all lines
  store intersection points in array
  K-d tree neighbourhood
  select neighbouring points within 2 pixels
  select cluster with maximum intersection points
  mean of selected intersection points

Next, the azimuths of the lines are calculated, and the maximum azimuth is selected based on the histogram of the azimuths. Subsequently, those LSD lines within the angle threshold of 15° from the maximum azimuth direction are selected. A random sample consensus (RANSAC) algorithm is applied to the selected lines to fit the optimal line and to get rid of falsely detected parallel lines. To extract lines more accurately, the Huber-robust line fitting algorithm in Kaehler and Bradski[35] with specified buffer width from each side of the fitted RANSAC line is applied. The neighbouring intersection points within 2 px from each intersection point are selected by means of the *K-d* tree neighbourhood algorithm. Finally, the maximum intersection point cluster is selected and their mean yields the final intersection point (Algorithm 1; see Omidalizarandi et al.[26] for further information).

Calibration of the optical measurement system

The displacement time series for the captured images of the Leica MS50 is produced by subtracting the extracted target centroids at different epochs of time. Subsequently, the calculated displacements in the pixel unit are converted to the more meaningful metric unit by means of the calibration parameters. The calibration consists of an internal calibration of the telescope camera of the Leica MS50 (regarding focal length, principal point and radial and tangential distortions) and an internal calibration of the error sources of the Leica MS50 measurements (including the zero offset for distance measurements or horizontal collimation error, vertical index error, tilting axis error and compensator index error of angular measurements). As mentioned previously, the FOV is reduced by $8 \times$ digital zoom to capture a small central portion of the image captured with $1 \times$ zoom (see Figure 2). Subsequently, it has very small impact of the aforementioned camera calibration parameters. However, as described in Ehrhart and Lienhart,[16] these camera calibration parameters can be neglected due to the relative calculation of the displacement for the sequences of the video frames. However, a proper design of the target pattern, as proposed in Omidalizarandi et al.[26] (see Figure 3), may eliminate the influences of the aforementioned small distortions using the redundant line features and extract them in the robust and reliable procedure. However, in this research, we treat the remaining systematic errors, such as calibration parameters, as coloured noise and separate them from the white noise based on an AR model of the coloured measurement noise (further discussion about this is in the next section). The displacement time series of the Leica MS50 is compared with the PSVC and the LT datasets to give an impression of the accuracy of the detected target centroid and to visually demonstrate the previous statements concerning the neglecting of the remaining calibration parameters. Since the output of the PSVC is the acceleration, it can be converted to the displacement in metric units based on equation (2)

$$\delta_{z_i} = \frac{a_{z_i} - \overline{a_z}}{(2\pi f)^2} \qquad (2)$$

where δ_{z_i} is the calculated displacement in the Z direction (mm), a_{z_i} is the acceleration in the Z direction (m/s^2), $\overline{a_z}$ is the average of the acceleration data within the specified period of time (m/s^2) and f is the frequency (Hz).

As we can see from Figure 4, the differences in the amplitudes for all three sensors are at a sub-millimetre-level accuracy. However, the time synchronization is still a challenge and needs to be performed precisely. In this work, the time synchronization is performed merely for the controlled excitation in the laboratory environment by changing the frequency and amplitude of the PSVC and by fitting the time series of all three sensors at a point of change. Subsequently, as we can also see
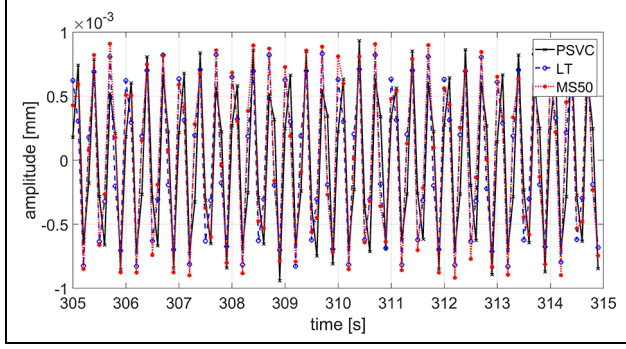
Figure 4. Displacement time series for the PSVC, the LT and the Leica MS50 at 3 Hz and distance of 7.52 m.
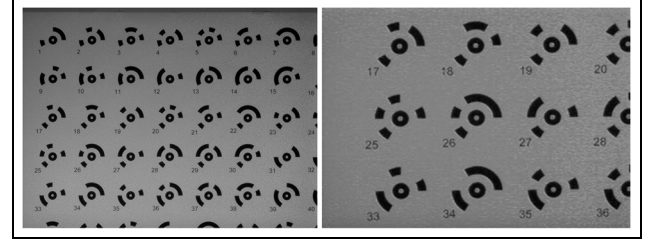

Figure 5. The coded target images captured using the telescope camera of the Leica MS50 at distances of 30.39 m (left) and 14.68 m (right).

$$r_c = \frac{\min(d_w, d_h)}{2} \text{ (px)} \qquad (5)$$

where w is the image width (px), h is the image height (px), d_w is the circle diameter in the horizontal direction (px), d_h is the circle diameter in the vertical direction (px), r_c is the circle radius (px) and FOV_h and FOV_v are the horizontal and vertical FOVs, which according to the user manual of the sensor equal 1.3° and 1.0°, respectively.

Next, the ellipse fitting in a least-squares sense is applied to the concentric edge contours with detected Hough circles. In order to assign a unique ID to each target and to detect coded targets, template matching is applied by their comparisons with the designed coded targets. Finally, the vertical angular conversion factor is calculated based on the equations

$$\beta = \cos^{-1}\left(\frac{\delta_y}{\delta_{xy}}\right) \text{ (rad)} \qquad (6)$$

$$psy = \frac{\delta_{XYZ} \cdot \cos(\beta)}{\delta_y} \text{ (mm/px)} \qquad (7)$$
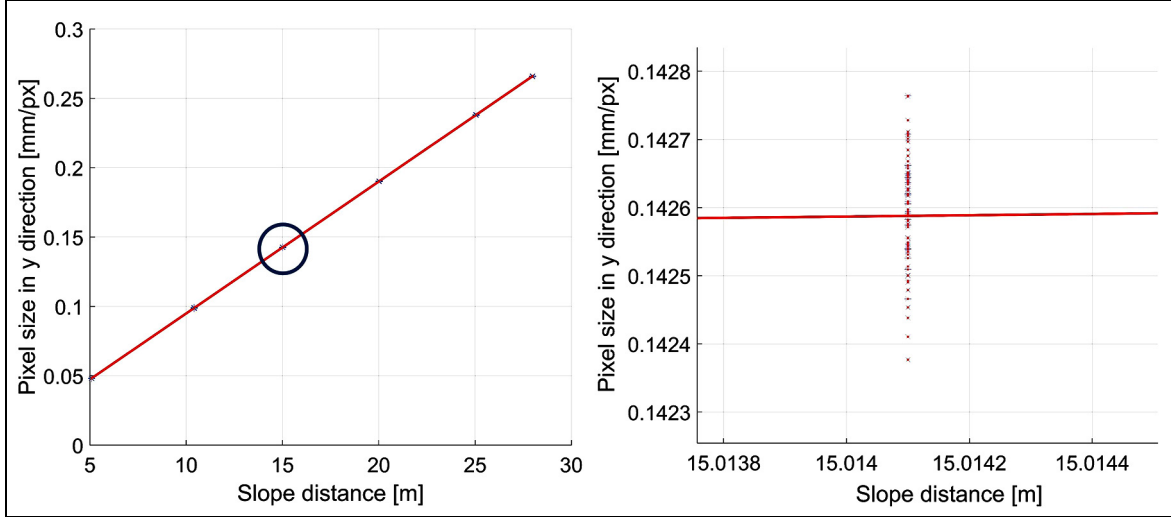
$$p = polyfit(psy, S) \qquad (8)$$

$$p_v = polyval(p, S) \qquad (9)$$

$$F\hat{O}V_v = \frac{p_v \cdot h}{S} \text{ (rad)} \qquad (10)$$

$$\alpha_v = \frac{F\hat{O}V_v}{h} \cdot \frac{180 \cdot 3600}{\pi} \text{ (''/px)} \qquad (11)$$

where δ_y is the difference between the target centroid in y direction (px), δ_{xy} is the difference between the target centroid in both x and y directions (px), δ_{XYZ} is the difference between the target centroid in X, Y and Z directions (m), psy is the pixel size in the y direction, p is the coefficient of the best-fitting polynomial of degree 1 in a least-squares sense, S is the slope distance (m), p_v is the value of the derived polynomial at specified slope distances, $F\hat{O}V_v$ is the calculated vertical FOV (rad), h is the height of the image (px) and α_v is the vertical angular conversion factor (''/px).

in Figure 4, the synchronization was not achieved perfectly. However, time synchronization is not in the focus of this research and will be investigated as part of future research.

The vertical angular conversion factor of the telescope camera of the Leica MS50 is calibrated to convert pixel quantities to metric quantities. The calibration is started by designing a coded target pattern in the software AutoCAD 2016 in the two paper sizes *A*2 and *A*4 with fixed coded target distances of 0.09 and 0.0335 m. As previously mentioned, the telescope camera of the Leica MS50 has a small FOV and cannot cover the entire target pattern at different distances (see Figure 5). Subsequently, the coded targets seem to be more advantageous compared to the chessboard pattern, since both the target centroid coordinates and the ID (i.e. target identification number) are obtained simultaneously. The highly accurate 3D object coordinates of the targets are obtained by taking multiple photos with a high-resolution camera from different viewing angles and by solving the space resection bundle adjustment in the iterative procedure.

The next step is to extract the target centroids of the coded target pattern images captured at different distances up to approximately 30 m and to assign them the unique IDs. To extract a target centroid, median blur and bilateral filtering are first applied to reduce the noise and to preserve the sharp edges of the images. The canny edge detector[36] is then applied to extract the edges. The inner circular part of the coded target is extracted by means of the Hough circle transform (HCT).[35] To find the circles based on the HCT, the circle radius is approximately calculated using equations (1) and (3)–(5)[26]

$$d_w = \frac{w \cdot FOV}{FOV_h \cdot \left(\frac{\pi}{180}\right)} \text{ (px)} \qquad (3)$$

$$d_h = \frac{h \cdot FOV}{FOV_v \cdot \left(\frac{\pi}{180}\right)} \text{ (px)} \qquad (4)$$

Figure 6. Depiction of the pixel sizes of the captured images in the direction of *y* axis with respect to the slope distances (left) and the magnification of the area highlighted by the circle (right).

The MATLAB functions *polyfit* and *polyval* are used, respectively, to fit a best line to the *psy* calculated for different slope distances and to evaluate it at specified distances. Figure 6 (left) depicts the *psy* values calculated for different slope distances, and Figure 6 (right) zooms in around the slope distance of 15 m.

The value 1.9583 (″/px) is obtained from the evaluation of the previous equations for calculation of the α_v, which is very close to the value 1.9632 (″/px) given in Ehrhart and Lienhart.[16] Furthermore, the calculated FOV_v is approximately 1.04442°, a value which is slightly different from the value given in the user manual of the sensor.

Displacement and vibration analysis

The discrete Fourier transform (DFT) is typically applied to estimate the amplitude and frequency of oscillating objects, such as bridge structures. It might achieve reasonable results while the measurements are less contaminated with the coloured noise. To tackle this problem and to estimate the amplitude and frequency even in the case of high coloured measurement noise, we proposed a robust and consistent procedure which can be extended and used for any type of measurement, particularly the vision-based measurement system, to obtain the highly accurate results. In this research, we utilize the captured video frames from the telescope camera of the Leica MS50 for displacement and vibration analysis of a footbridge structure.

We use a simple harmonic motion to perform the displacement measurements, which means that the acceleration measurement is directly proportional to its displacement from the equilibrium position. In addition, the acceleration is directed towards the equilibrium position.[37] Subsequently, the extracted target centroids from video streams of the Leica MS50 or the 3D coordinates from the LT are always averaged over a specified period of time to define the equilibrium position. We can calculate the displacement and acceleration for Leica MS50 measurements using the equations

$$\delta_{y_i} = \alpha S(y_i - \bar{y}) \tag{12}$$

$$\delta_{a_i} = \alpha S(y_i - \bar{y})(2\pi f)^2 \tag{13}$$

where S is the slope distance (m), α is the vertical angular conversion factor (″/px), y_i is the extracted target centroid at epoch i (px), $\bar{y}$ is the average of the extracted target centroid within the specified period of time (px) and f is the frequency (Hz).

To compare the Leica MS50 and the LT, we can calculate the displacements for the Leica MS50 measurements based on equation (12) and then input the displacements to equation (15) to calculate the amplitude (mm) and the frequency (Hz), respectively. Concerning the comparison of the Leica MS50 and the PSVC, we note that the output of the latter consists of acceleration measurements; we can calculate accelerations for the Leica MS50 measurements based on equation (13) and then use equation (15) to calculate the amplitude (m/s²) and the frequency (Hz). However, it is also possible to calculate displacements from the acceleration measurements via double integration from equation (15).

To estimate the frequency, merely pixel differences are sufficient to derive reasonable results due to the linearity property

$$F\{cf(t)\} = cF\{f(t)\} \tag{14}$$

of the Fourier transform.[16,38] Despite this possibility, we performed both displacement and vibration analysis in metric unit measurements.

We modelled the given vibration measurements $\ell_1, \ldots, \ell_n$ by means of a sum of sinusoids and additive random deviations $e_1, \ldots, e_n$, as proposed by Kargoll et al.,[39] that is

$$\ell_t = \frac{a_0}{2} + \sum_{j=1}^{M} \left[a_j \cos(2\pi f_j x_t) + b_j \sin(2\pi f_j x_t) \right] + e_t \tag{15}$$

The frequencies $f_1, \ldots, f_M$ and the coefficients a_0, $a_1, \ldots, a_M$ and $b_1, \ldots, b_M$ are treated as unknown parameters. Collecting these unknowns within the vector $\boldsymbol{\xi}$, we can write the preceding non-linear observation equations in the form $\ell_t = h_t(\boldsymbol{\xi}) + e_t$. To take coloured measurement noise into account, we assume the random deviations to be autocorrelated through a covariance-stationary AR process

$$e_t = \alpha_1 e_{t-1} + \cdots + \alpha_p e_{t-p} + u_t \tag{16}$$

in which the coefficients $\boldsymbol{\alpha}^T = [\alpha_1, \ldots, \alpha_p]$ are also considered as unknown parameters. Since we expect numerous outliers of different magnitudes to be present in the data, the white noise components $u_1, \ldots, u_n$ are assumed to follow the centred and scaled t-distribution $t_\nu(0, \sigma^2)$ independently with an unknown degree of freedom ν and with unknown scale factor σ^2. We, thus, have a one-dimensional version of the generic observation model in section 'Sensor specifications and measurement systems' in Alkhatib et al.[30] Estimation of the model parameters $\boldsymbol{\theta}^T = [\boldsymbol{\xi}^T, \boldsymbol{\alpha}^T, \sigma^2, \nu]^T$ can, thus, be carried out efficiently by means of a one-dimensional version of the GEM algorithm given in section 'Passive target centroid detection' in Alkhatib et al.[30] To develop this algorithm, the observation and AR equations are inverted into

$$u_t = e_t - \alpha_1 e_{t-1} - \cdots - \alpha_p e_{t-p} = \alpha(L)(\ell_t - h_t(\boldsymbol{\xi})) \tag{17}$$

where the lag operator $L^j e_t := e_{t-j}$ and the lag polynomial $\alpha(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p$ are used as convenient notations. The required initial values e_0, e_{-1} and so on are set equal to zero for simplicity.

Within iteration step i of this algorithm, the E-step consists of the adjustment of the observation weights

$$w_t^{(i)} = \frac{\nu^{(i)} + 1}{\nu^{(i)} + \left(u_t^{(i)}/\sigma^{(i)} \right)^2} \tag{18}$$

which depend on currently available initial or estimated parameter values $\boldsymbol{\xi}^{(i)}$, $\boldsymbol{\alpha}^{(i)}$, $\sigma^{(i)}$ and $\nu^{(i)}$. The individual weights then give rise to the diagonal weight matrix $\mathbf{W}^{(i)}$ used within the subsequent M-step. Initially, we may use the unit weight matrix $\mathbf{W}^{(0)} = \mathbf{I}_n$, the vanishing AR process $\boldsymbol{\alpha}^{(0)} = \mathbf{0}_{[p \times 1]}$, the identity scale factor $\sigma^{(i)} = 1$ and the degree of freedom $\nu^{(0)} = 30$, and these choices correspond to the initial assumption of approximately standard-normal and uncorrelated random deviations.

The M-step can be carried out by solving the four parameter groups individually. First, the parameters $\boldsymbol{\xi}$ are determined by solving the linearised normal equations

$$\boldsymbol{\Delta \xi}^{(i+1)} = \left((\overline{\mathbf{A}}^{(i)})^T \mathbf{W}^{(i)} \overline{\mathbf{A}}^{(i)} \right)^{-1} (\overline{\mathbf{A}}^{(i)})^T \mathbf{W}^{(i)} \overline{\boldsymbol{\Delta \ell}}^{(i)} \tag{19}$$

with reduced observations $\Delta \ell_t^{(i)} = \ell_t - h_t(\boldsymbol{\xi}^{(i)})$, decorrelation filtered reduced observations $\overline{\Delta \ell_t^{(i)}} := \alpha^{(i)}(L)\Delta \ell_t$, Jacobi matrix components $A_{t,k}^{(i)} = \partial h_t(\boldsymbol{\xi}^{(i)})/\partial \xi_k$ and decorrelation filtered Jacobi matrix components $\overline{A_{t,k}^{(i)}} := \alpha^{(i)}(L)A_{t,k}$. By virtue of the functional relationship (equation (15)), the derivatives occurring read

$$\frac{\partial h_t(\boldsymbol{\xi}^{(i)})}{\partial a_0} = \frac{1}{2}, \quad \frac{\partial h_t(\boldsymbol{\xi}^{(i)})}{\partial a_j} = \cos\left(2\pi f_j^{(i)} x_t\right)$$

$$\frac{\partial h_t(\boldsymbol{\xi}^{(i)})}{\partial b_j} = \sin\left(2\pi f_j^{(i)} x_t\right)$$

$$\frac{\partial h_t(\boldsymbol{\xi}^{(i)})}{\partial f_j} = 2\pi t \left[b_j^{(i)} \cos\left(2\pi f_j^{(i)} x_t\right) - a_j^{(i)} \sin\left(2\pi f_j^{(i)} x_t\right) \right]$$

A Gauss–Newton step with step size $\gamma \in (0, 1]$ gives

$$\boldsymbol{\xi}^{(i+1)} = \boldsymbol{\xi}^{(i)} + \gamma \boldsymbol{\Delta \xi}^{(i+1)} \tag{20}$$

and this solution yields the estimated coloured noise residuals $e_t^{(i+1)} = \ell_t - h_t(\boldsymbol{\xi}^{(i+1)})$. Based on these residuals, we assemble the matrix

$$\mathbf{E}^{(i+1)} = \begin{bmatrix} e_0^{(i+1)} & \cdots & e_{1-p}^{(i+1)} \\ \vdots & & \vdots \\ e_{n-1}^{(i+1)} & \cdots & e_{n-p}^{(i+1)} \end{bmatrix} \tag{21}$$

and then compute the solution of the normal equations with respect to the AR coefficients

$$\boldsymbol{\alpha}^{(i+1)} = \left((\mathbf{E}^{(i+1)})^T \mathbf{W}^{(i)} \mathbf{E}^{(i+1)} \right)^{-1} (\mathbf{E}^{(i+1)})^T \mathbf{W}^{(i)} \mathbf{e}^{(i+1)} \tag{22}$$

Here, we need to check whether all roots of $\alpha^{(i+1)}(z) = 0$ are located within the unit circle; if this is not true, we mirror all roots with a magnitude larger than 1 into the unit circle in order for the estimated AR process to be invertible. Applying the inversion (equation (17)) to this process, the white noise residuals

result from the coloured noise residuals through $u_t^{(i+1)} = \alpha_t^{(i+1)}(L)e_t^{(i+1)}$. The scale factor is now computed as the weighted sum of squared white noise residuals divided by the number of observations, that is

$$(\sigma^2)^{(i+1)} = \frac{1}{n}\sum_{t=1}^{n} w_t^{(i)}\left(u_t^{(i+1)}\right)^2 \tag{23}$$

Finally, to estimate the degree of freedom of the underlying *t*-distribution, we determine the zero of the equation

$$\begin{aligned}0 = {} & \log \nu^{(i+1)} + 1 - \psi\left(\tfrac{\nu^{(i+1)}}{2}\right) + \psi\left(\tfrac{\nu^{(i+1)}+1}{2}\right) \\ & - \log\left(\nu^{(i+1)}+1\right) + \tfrac{1}{n}\sum_{t=1}^{n}\left(\log w_t^{(i+1)} - w_t^{(i+1)}\right)\end{aligned} \tag{24}$$

where the weights $w_t^{(i+1)}$ are defined as in equation (18) by substituting $u_t^{(i+1)}$, $\sigma^{(i)}$ and the variable $\nu^{(i)}$, and where ψ denotes the digamma function. More details on the derivation and the implementation of this algorithm can be found in Alkhatib et al.[30]

Experiments and results

We performed accurate displacement and vibration analysis using video streams from the telescope camera of a Leica MS50 with a practical sampling frequency of 10 Hz. Experiments were performed for two case studies under (1) a controlled laboratory environment and (2) an uncontrolled real-world situation observing a footbridge structure using the telescope camera of Leica MS50. Furthermore, an LT and a PSVC were used as two highly accurate reference sensors with a sampling frequency of 200 Hz for the validation purposes. Alternatively, the calculated natural frequencies of the footbridge structure based on the FEM were utilized for a validation.

The primary step to perform a displacement and vibration analysis based on the video frames of the Leica MS50 was to select an optimal passive target pattern and to extract its centroid with high accuracy at different epochs of time. Next, the vertical angular conversion factor of the telescope camera of the Leica MS50 was calibrated, which allows us to convert derived displacements from the pixel unit to the metric unit. In addition, the Fourier series (equation (15)) as a linear regression model and an AR process (equation (16)) as a coloured noise model were employed to estimate amplitudes and frequencies with high accuracy, assuming the white-noise components to follow a scaled *t*-distribution with an unknown scale factor and unknown degree of freedom. To estimate the model parameters by means of the GEM algorithm described in the preceding section, the number M of Fourier frequencies and the model order p of AR process were

specified beforehand. We determined the initial values $f_1^{(0)}, \ldots, f_M^{(0)}$ for the unknown frequencies based on notable maximum amplitudes within the DFT of the data. In addition, the model order of the AR process was set to 25 throughout the entire procedure. To apply the GEM algorithm, all three datasets were divided to the segments of $n = 1000$ consecutive measurements, spanning approximately 5 s for the LT and PSVC measurements and spanning 100 s for the Leica MS50 measurements.

Example based on the shaker vibration calibrator

The controlled excitations were performed at the laboratory of the Geodetic Institute Hannover (GIH) of the Leibniz Universität Hannover (LUH; see Figure 7, right) by means of the Leica MS50, LT, PSVC and IMU Brick 2.0 (which constitutes a low-cost accelerometer). The analysis of the IMU Brick 2.0 measurements is beyond the scope of this article and will be carried in our future research. As can be seen from Figure 7 (left), the optimal passive target pattern as proposed by Omidalizarandi et al.[26] oscillates as part of the PSVC and is simultaneously measured through video streaming by means of the telescope camera of the Leica MS50. The PSVC contains a PCB ICP quartz reference accelerometer, which outputs highly accurate acceleration data. Specifically, the oscillation frequencies of 2, 3 and 4 Hz with an amplitude of 0.3 m/s^2 were adjusted throughout two sensitive dials. Each frequency was measured for a about 7 min by all four sensors. However, since the sampling frequency of the Leica MS50 is only 10 Hz in practice, the frequencies of

Figure 7. Vibration analysis of a controlled excitation based on acceleration measurements from the PSVC 9210D, video streams from the telescope camera of the Leica MS50, 3D coordinates from the LT Leica AT960-LR and acceleration measurements from the IMU Brick 2.0.
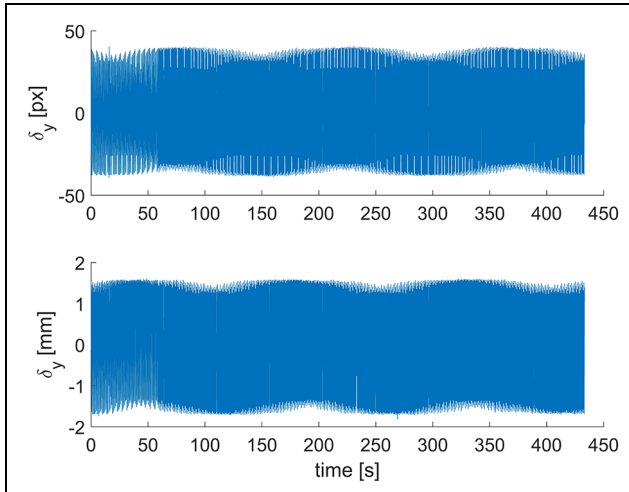
Figure 8. Displacement time series of the extracted target centroid for the telescope camera images of the Leica MS50 at 2 Hz and a distance of 5.3616 m.

Figure 9. Typical discrete Fourier transform of one segment of the Leica MS50 dataset at distance of 5.3616 m, showing the main amplitude at 2 Hz.

higher than 5 Hz could not be captured by the Leica MS50 (in view of the Nyquist sampling theorem). To perform measurements with the LT, a Leica red-ring reflector (RRR) 0.5 in ball with a radius of 6.35 mm ± 0.0025 mm and an acceptance angle $\leq \pm 30°$ was mounted on the small platform which vibrates simultaneously with aforementioned passive target.

Figure 8 shows the displacement time series of the extracted target centroid with respect to the mean of the extracted centroids throughout time in both millimetre and pixel units at 2 Hz frequency and a slope distance of 5.3616 m. In addition, the first 20 s of the Leica MS50 data and the first 5 s of the LT and PSVC data were discarded as transient oscillations.

Figure 9 depicts the DFT of the video streams from the Leica MS50 at a distance of 5.3616 m, where the frequency induced by the PSVC was 2 Hz. This frequency is clearly associated with the maximum amplitude. Figure 10 shows the target centroid extracted from the Leica MS50 video streams alongside the adjusted Fourier model at 2 Hz for a 5 s time section.

Figure 11 shows the estimated coloured noise residuals and the decorrelated residuals of the Leica MS50 dataset, resulting from the filtering (equation (17)) of the former residuals by means of the inverted estimated AR model. Figure 12 shows the adequacy of the estimated AR coloured noise models in the light of an accepted (periodogram-based) white noise.[39]

The DFT is shown in Figure 13, which reveals two main amplitudes at 1.25 and 2.5 Hz to shed further light on the impact of the image motion error (see Figure 14) on the estimation of the frequency. In addition, Figure 15 shows a higher coloured noise level in comparison to Figure 11, which proves the existence of
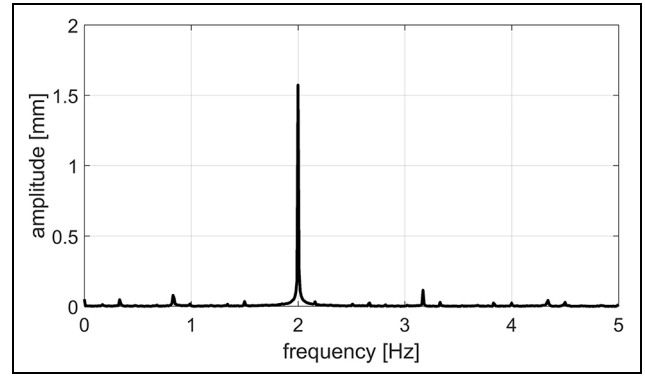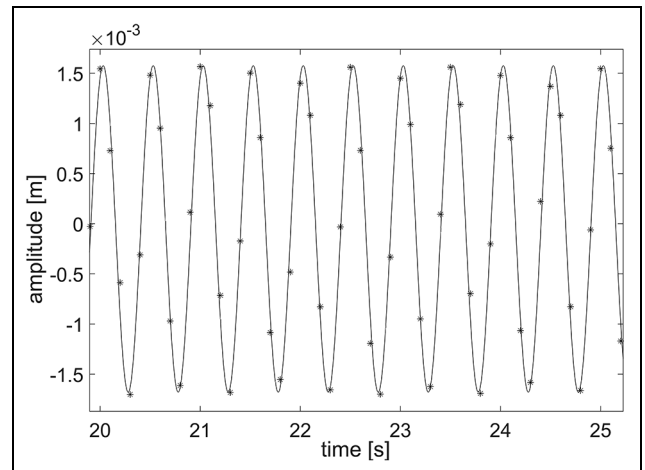
Figure 10. Typical section of the Fourier model (solid line) fitted to the given measurements of the Leica MS50 dataset (stars) at 2 Hz within 5 s.
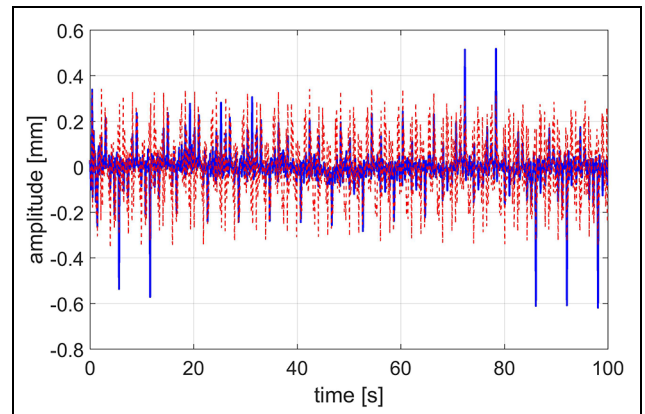
Figure 11. Typical segment showing the estimated coloured noise residuals and the decorrelated residuals of the Leica MS50 dataset at 2 Hz and a distance of 5.3616 m.
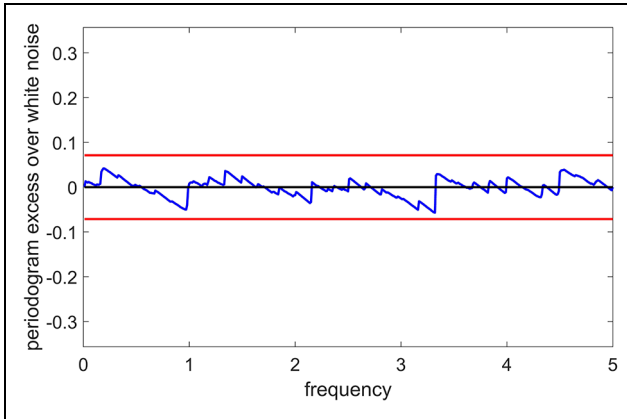
Figure 12. Excess of the estimated periodogram of the decorrelated (i.e. estimated white noise) residuals of the Leica MS50 dataset at 2 Hz and a distance of 5.3616 m for the AR(25) model (jagged line) with respect to the theoretical white noise periodogram (horizontal centred line) and 99% significance bounds (horizontal bounded lines).

Figure 13. Typical discrete Fourier transform of one segment of the Leica MS50 dataset at a distance of 22.6635 m, showing two main amplitudes at 1.25 and 2.5 Hz.

the image motion errors throughout this time interval of the experiment.

Table 1 summarizes the statistics of the displacement and vibration analysis for all three sensors. In most cases, the frequencies and amplitudes estimated from the Leica MS50 measurements are very close to those resulting from the two highly accurate reference sensors (LT and PSVC). The estimated degrees of freedom of the t-distribution underlie the white noise components. Concerning the LT and the PSVC measurements, these estimates are roughly between 14 and 60, indicating a rather close approximation of a normal distribution. By contrast, the estimated degrees of freedom regarding the Leica MS50 measurements are in the range of 2–4.5, for which values the t-distribution has substantial tails; we thus found a large number of outliers in the measurement noise of that sensor.

Furthermore, the results show that the highest white noise test acceptance rate (100%) was obtained for the PSVC measurements, for which an AR order of $p = 25$ was chosen. Using the same model order, the LT measurements also produced relatively high acceptance rates in comparison to the PSVC data. However, the acceptance rates regarding the Leica MS50 measurements fluctuate between 25% and 75%, so that the adjusted coloured noise model is clearly inadequate for a number of segments analysed. On one hand, this finding could be related to the image motion error (see Figure 14) derived from a weak PC performance or delay in the data transmission procedure from the Leica MS50 to the PC. On the other hand, it could be related to a minor shaking of the Leica MS50 throughout the measurements, and this phenomenon can be taken into account in our future work by continuously reading the
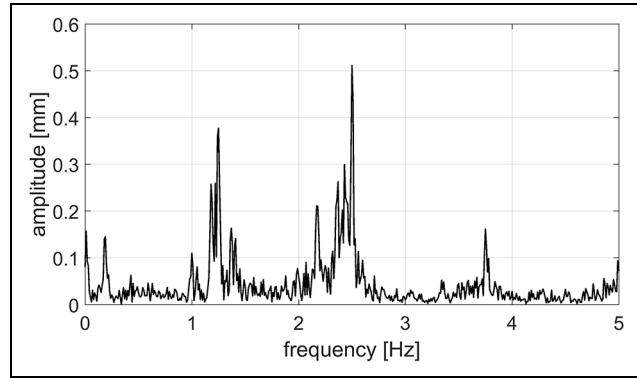
tilting axis error of the Leica MS50 using the GeoCOM interface as described in Lienhart and colleagues.[3,16] To improve the coloured noise models and the resulting performance of the white noise test, it might be beneficial in future experiments to increase the measurement time to obtain more redundant data and to be able to increase the AR model order. Moreover, the absolute deviation of the *Mode* of estimated amplitudes of the sensors are listed in Table 1 and compared to those from the PSVC with an AR order of $p = 25$. As we expected, the absolute deviations of the two reference sensors of LT and the PSVC have minor differences, which are significantly smaller than those of the Leica MS50.

Table 2 summarizes the statistics of the displacement and vibration analysis for all three sensors without AR processing to give an impression about strength of the developed algorithm. In addition, the degree of freedom was fixed to 120, which stands for the t-distribution, approximating the normal distribution as described in Abramowitz and Stegun[40] and Koch.[41] As we expected, the absolute deviations from the PSVC without the AR processing are not significant and has less coloured measurement noise, as the estimated degree of freedoms indicate approximately the normal distribution (see Table 1). In addition, the absolute deviations for the LT measurement without the AR processing are slightly larger than those including AR processing. However, the absolute deviations for the Leica MS50 without the AR processing are mostly and significantly larger than those included in the AR processing. In addition, as the estimated degrees of freedom for the Leica MS50 data are represented by a range of 2–4.5 (see Table 1), it proves the existence of numerous outliers in the dataset. Subsequently, by ignoring the AR processing within the robust estimation procedure developed for the Leica MS50, the results in some cases do not prove to be reliable or accurate enough.

Table 1. Statistics of the displacement and vibration analysis for the Leica MS50, LT and PSVC measurements with an AR order of $p = 25$.

Sensor	Mode ($\hat{f}$, Hz)	Time (s)	NS	S (m)	Max ($\hat{L}$, mm)	Mode ($\hat{L}$, mm)	Min ($\hat{L}$, mm)	Mode ($\hat{v}$)	WNT (%)	Δ_d (mm)
LT	2.000067	5	85	5.3616	1.7476	1.7358	1.7067	28.467	94.11	0.0327
	3.000077	5	83	5.3616	0.8264	0.8203	0.8002	53.229	78.31	0.0011
	4.000103	5	87	5.3616	0.4616	0.4598	0.4567	54.345	66.66	0.0032
	2.000077	5	80	11.9425	1.7708	1.7565	1.7478	59.547	96.25	0.0269
	3.000071	5	104	11.9425	0.8221	0.8186	0.8148	51.202	82.69	0.0008
	4.000105	5	83	11.9425	0.4608	0.4578	0.4545	50.171	75.90	0.0057
	2.000042	5	85	16.8771	1.7378	1.7256	1.7066	50.212	96.47	0.0310
	3.000152	5	82	16.8771	0.8271	0.8151	0.8025	28.257	75.60	0.0044
	4.000112	5	127	16.8771	0.4606	0.4574	0.4497	38.732	82.67	0.0061
	2.000027	5	99	22.6635	1.7455	1.7231	1.7094	26.180	90.90	0.0321
	3.000061	5	86	22.6635	0.8333	0.8246	0.7994	14.272	87.21	0.0044
	4.000070	5	84	22.6635	0.4694	0.4665	0.4518	18.819	29.76	0.0035
PSVC	2.000106	5	95	5.3616	1.7227	1.7031	1.675	44.371	100	0.0
	3.000221	5	99	5.3616	0.8300	0.8192	0.8006	54.612	100	0.0
	4.000146	5	101	5.3616	0.4665	0.4630	0.4595	53.664	100	0.0
	2.000142	5	91	11.9425	1.7514	1.7296	1.7041	47.149	100	0.0
	3.000186	5	93	11.9425	0.8262	0.8194	0.8104	51.829	100	0.0
	4.000299	5	93	11.9425	0.4667	0.4635	0.4569	50.199	100	0.0
	2.000099	5	97	16.8771	1.7166	1.6946	1.6755	40.411	100	0.0
	3.000128	5	97	16.8771	0.8256	0.8195	0.8043	49.681	100	0.0
	4.000220	5	137	16.8771	0.4665	0.4635	0.4593	42.223	100	0.0
	2.000075	5	121	22.6635	1.7111	1.6910	1.6721	45.319	100	0.0
	3.000178	5	113	22.6635	0.8258	0.8202	0.8118	54.992	100	0.0
	4.000215	5	133	22.6635	0.4661	0.4630	0.4602	48.468	100	0.0
MS50	2.001359	100	4	5.3616	1.6256	1.6181	0.4063	3.8933	75	0.0850
	3.002026	100	4	5.3616	0.7496	0.7473	0.7424	4.4704	75	0.0719
	4.002689	100	4	5.3616	0.4849	0.4119	0.4091	3.4502	25	0.0511
	2.001327	100	4	11.9425	1.7250	1.7184	1.7147	4.4514	75	0.0112
	3.003544	100	4	11.9425	0.7888	0.7842	0.6814	4.2560	25	0.0352
	4.002680	100	4	11.9425	0.4339	0.4322	0.4287	1.9757	25	0.0313
	2.002867	100	4	16.8771	1.6986	1.5932	0.2670	4.4179	50	0.1014
	3.002015	100	4	16.8771	0.8293	0.7925	0.7678	3.4979	50	0.0270
	4.002692	100	3	16.8771	0.4352	0.4347	0.4225	3.1020	33.33	0.0288
	2.001887	100	4	22.6635	1.7356	1.4911	0.1650	1.9504	25	0.1999
	3.002021	100	4	22.6635	0.7808	0.7591	0.7536	3.3790	75	0.0611
	4.002694	100	4	22.6635	0.4240	0.4209	0.3817	2.8920	25	0.0421

LT: laser tracker; PSVC: portable shaker vibration calibrator; AR: autoregressive.
$\hat{f}$ is the estimated frequency, NS the number of segments analysed, S the slope distance, $\hat{L}$ the estimated amplitude or displacement, $\hat{v}$ the estimated degree of freedom, WNT the acceptance rate of the white noise test applied to the NS segments analysed and Δ_d the absolute deviation of the mode ($\hat{L}$) of the sensors compared to those from PSVC.

Example based on real application of a footbridge structure

An uncontrolled excitation of a footbridge structure with a length of 27.051 m and a width of 2.72 m close to the GIH (see Figure 16) was measured using the Leica MS50 and LT. The measurements were carried out for the first quarter and the middle of the footbridge structure (marked by the circles in Figure 16). Alternatively, the known natural frequencies of the footbridge structure were utilized for a validation. They were calculated based on the FEM analysis of a design model of the footbridge structure, which was carried out by the Institute of Concrete Construction of the LUH. According to that, the vertical natural frequencies of the footbridge structure are 3.642 and 13.294 Hz, the longitudinal and the lateral natural frequencies are 2.295 and 7.053 Hz, and the torsional natural frequencies are 3.759 and 11.828 Hz, respectively. As we can see in Figure 16, the Leica MS50 and the LT are located at the footpath close to the side of the footbridge structure. Regarding the natural frequencies, we could only detect the vertical natural frequency of 3.642 Hz and could not detect another one with the value of 13.294 Hz due to the low practical sampling frequency of 10 Hz of the Leica MS50 and in view of Nyquist sampling theorem. On the other hand, it might be necessary to set-up the Leica MS50 in a place, where

Table 2. Statistics of the displacement and vibration analysis for the Leica MS50, LT and PSVC measurements without AR model and $\nu = 120$ the degree of freedom.

| Sensor	Mode ($\hat{f}$, Hz)	Time (s)	NS	S (m)	Max ($\hat{L}$, mm)	Mode ($\hat{L}$, mm)	Min ($\hat{L}$, mm)	Mode ($\hat{\nu}$)	WNT (%)	Δ_d (mm)
LT	2.000060	5	85	5.3616	1.7480	1.7363	1.7076	120	0	0.0332
	3.000081	5	83	5.3616	0.8263	0.8203	0.8003	120	0	0.0011
	4.000134	5	87	5.3616	0.4617	0.4597	0.4568	120	0	0.0033
	2.000075	5	80	11.9425	1.7707	1.7565	1.7481	120	0	0.0269
	3.000089	5	104	11.9425	0.8223	0.8187	0.8148	120	0	0.0007
	4.000113	5	83	11.9425	0.4606	0.4579	0.4546	120	0	0.0056
	2.000072	5	85	16.8771	1.7364	1.7247	1.7063	120	0	0.0301
	3.000158	5	82	16.8771	0.8312	0.8158	0.8014	120	0	0.0037
	4.000096	5	127	16.8771	0.4652	0.4575	0.4525	120	0	0.0060
	2.000060	5	99	22.6635	1.7452	1.7228	1.7097	120	0	0.0318
	3.000103	5	86	22.6635	0.8394	0.8251	0.8057	120	0	0.0049
	4.000111	5	84	22.6635	0.4732	0.4669	0.4618	120	0	0.0039
PSVC	2.000087	5	95	5.3616	1.7231	1.7032	1.6747	120	0	0.0001
	3.000249	5	99	5.3616	0.8298	0.8194	0.8006	120	0	0.0002
	4.000129	5	101	5.3616	0.4665	0.4630	0.4595	120	0	0.0000
	2.000204	5	91	11.9425	1.7503	1.7289	1.7040	120	0	0.0007
	3.000169	5	93	11.9425	0.8262	0.8193	0.8102	120	0	0.0001
	4.000279	5	93	11.9425	0.4667	0.4633	0.4568	120	0	0.0002
	2.000151	5	97	16.8771	1.7176	1.6941	1.6748	120	0	0.0005
	3.000116	5	97	16.8771	0.8253	0.8196	0.8042	120	0	0.0001
	4.000233	5	137	16.8771	0.4666	0.4635	0.4592	120	0	0.0000
	2.000024	5	121	22.6635	1.7115	1.6912	1.6743	120	0	0.0002
	3.000144	5	113	22.6635	0.8258	0.8204	0.8118	120	0	0.0002
	4.000229	5	133	22.6635	0.4662	0.4631	0.4602	120	0	0.0001
MS50	2.001361	100	4	5.3616	1.6205	1.6161	1.0136	120	0	0.0870
	3.002032	100	4	5.3616	0.7492	0.7462	0.7449	120	0	0.0730
	4.002711	100	4	5.3616	0.4067	0.4059	0.3928	120	0	0.0571
	2.001328	100	4	11.9425	1.7182	1.7125	1.7012	120	0	0.0171
	3.002494	100	4	11.9425	0.7857	0.7490	0.3325	120	0	0.0704
	4.002717	100	4	11.9425	0.4276	0.4261	0.4222	120	0	0.0374
	2.001526	100	4	16.8771	1.6960	1.5940	0.8262	120	0	0.1006
	3.002020	100	4	16.8771	0.7934	0.7840	0.7752	120	0	0.0355
	4.002711	100	3	16.8771	0.4325	0.3174	0.1332	120	0	0.1461
	2.001868	100	4	22.6635	1.0612	0.5376	0.4042	120	0	1.1534
	3.002025	100	4	22.6635	0.7619	0.7578	0.7422	120	0	0.0624
	4.002723	100	4	22.6635	0.4212	0.4178	0.4114	120	0	0.0452

LT: laser tracker; PSVC: portable shaker vibration calibrator; AR: autoregressive.

$\hat{f}$ is the estimated frequency, NS the number of segments analysed, S the slope distance, $\hat{L}$ the estimated amplitude or displacement, WNT the acceptance rate of the white noise test applied to the NS segments analysed and Δ_d the absolute deviation of the Mode ($\hat{L}$) of the sensors compared to those from PSVC in Table 1.

it can measure the passive targets perpendicularly to detect the longitudinal and lateral natural frequencies of the footbridge structure. However, this was not the case in our measurement campaign and we could not detect those natural frequencies either.

Figures 17 and 18 show the DFT results for the Leica MS50 measurement at Points A1 and A2, respectively. In addition, Figure 19 shows the DFT result at point A1 for the LT measurement as well. Since the time synchronization needed to be performed between the Leica MS50 and the LT measurements and was not the case in our measurement campaign, the one-to-one comparison between the results of the aforementioned sensors does not make sense. However, we can rely on

the FEM results of the design model of the footbridge structure for the validation in this case study. As can be seen from Figures 17 and 19, the DFT results at point A1 show the frequency of 4.07 Hz with the maximum amplitude of 0.07 mm, which was captured by the Leica MS50 measurements, whereas the frequency of 4.06 Hz with the maximum amplitude of 0.0497 mm was obtained for the LT measurements. However, the results from the proposed approach show the frequencies of 4.0768 and 4.0631 Hz with the maximum amplitudes of 0.0433 and 0.0462 mm for the Leica MS50 and LT measurements, respectively. However, the DFT result in Figure 18 at point A2 shows the frequencies of 3.65 and 4.1 Hz with the maximum amplitudes of 0.079
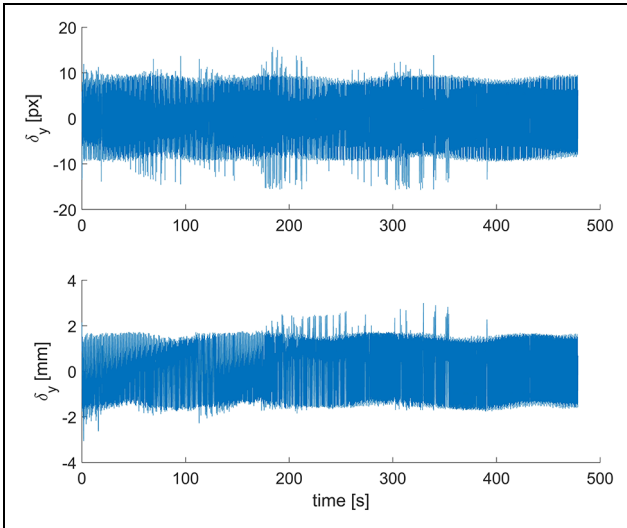
Figure 14. Displacements of the extracted target centroid for telescope camera images of Leica MS50 at 2 Hz and a distance of 22.6635 m.
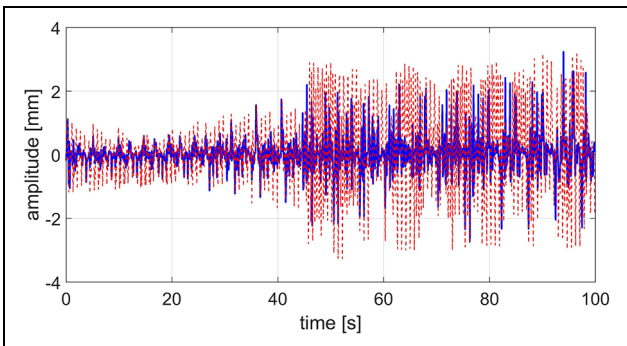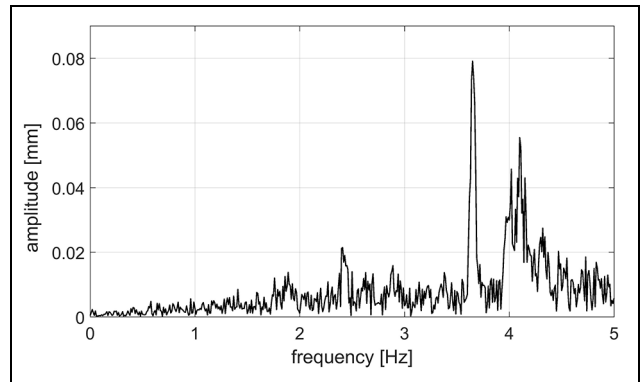

Figure 15. Typical segment showing the estimated coloured noise residuals (dashed line) and the estimated white noise residuals (solid line) of the Leica MS50 dataset at 2 Hz and a distance of 22.6635 m.
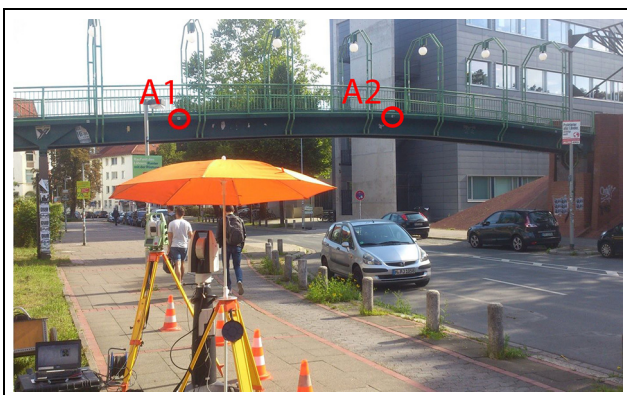

Figure 16. Vibration analysis of the footbridge structure using the telescope camera of the Leica MS50 to measure the passive targets attached to the side of the footbridge (areas highlighted by the circles) and using the LT to measure an attached corner cube reflector close to the aforementioned passive targets.
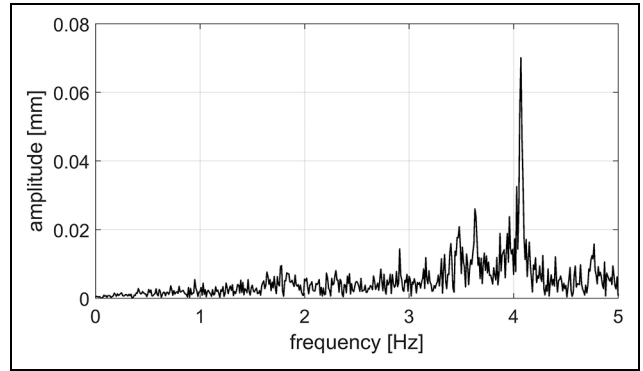
Figure 17. Point *A1*: Typical discrete Fourier transform of one segment of the Leica MS50 dataset at a distance of 16.5 m from footbridge structure, showing the main amplitude of 0.07 mm at a frequency of 4.07 Hz.


Figure 18. Point *A2*: Typical discrete Fourier transform of one segment of the Leica MS50 dataset at a distance of 20.01 m from footbridge structure, showing the main amplitudes of 0.079 and 0.055 mm at frequencies of 3.65 and 4.1 Hz, respectively.
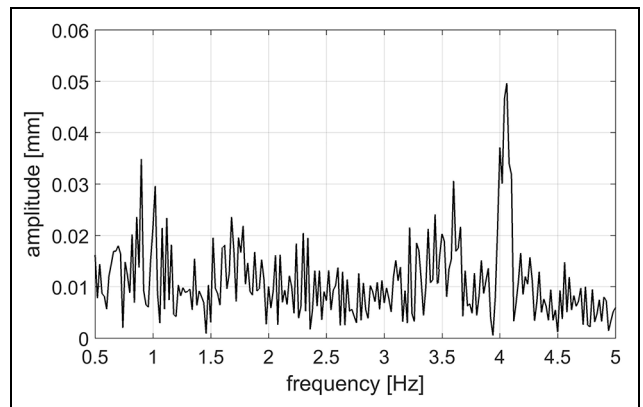
Figure 19. Point *A1*: Typical discrete Fourier transform of one segment of the LT dataset at a distance of 16.499 m from the footbridge structure, showing the main amplitude of 0.0497 mm at a frequency of 4.06 Hz.
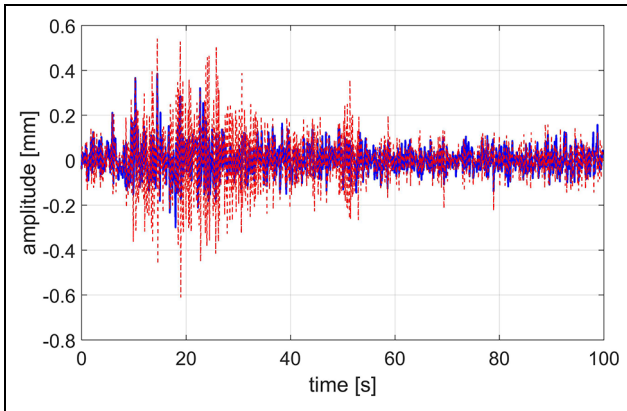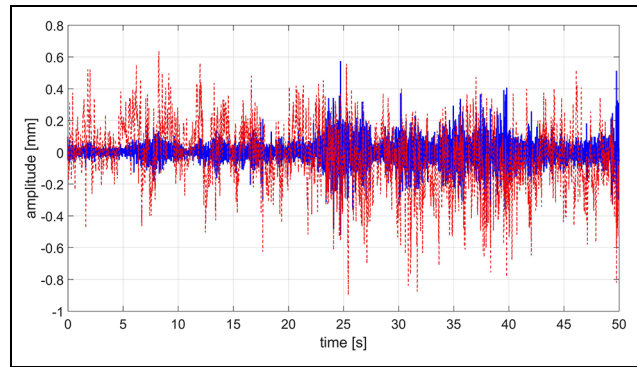
Figure 20. Point A1: Typical segment showing the estimated coloured noise residuals (dashed line) and the estimated white noise residuals (solid line) of the Leica MS50 dataset with the main amplitude at 4.07 Hz for the footbridge structure.
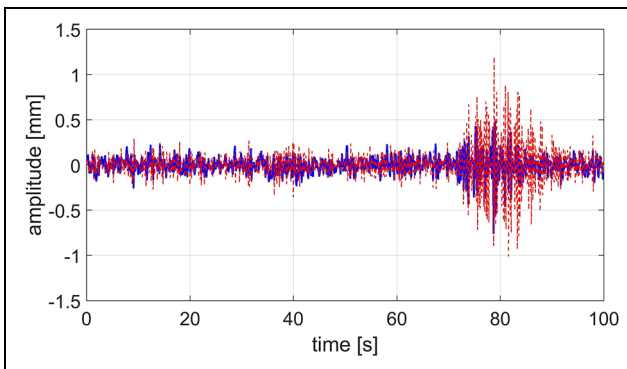
Figure 21. Point A2: Typical segment showing the estimated coloured noise residuals (dashed line) and the estimated white noise residuals (solid line) of the Leica MS50 dataset with the main amplitudes at frequencies of 3.65 and 4.1 Hz for the footbridge structure.

and 0.055 mm, respectively. However, the proposed approach shows the frequencies of 3.6451 and 4.0978 Hz with the maximum amplitudes of 0.051 and 0.0367 mm, respectively, which are comparable to those calculated from the FEM with the frequency of 3.642 Hz, and these findings demonstrate the correctness of our calculations. The higher frequency of approximately 4.07 Hz obtained might be due to either modulating the higher frequency of 13.294 Hz or an additional vibration produced by people passing across the footbridge structure and the latter not reaching a stable situation at the time of measurement.

In addition, Figures 20–22 illustrate the estimated coloured noise residuals and the decorrelated (i.e. estimated white noise) residuals for the Leica MS50 at the two points of A1 and A2 in addition to the LT measurements, respectively. We should mention that the LT

Figure 22. Point A1: Typical segment showing the estimated coloured noise residuals (dashed line) and the estimated white noise residuals (solid line) of the LT dataset with the main amplitude at 4.06 Hz for the footbridge structure.

measurements of the bridge structure were performed with a corner cube reflector of lesser quality than for measurements within the laboratory experiment. Consequently, a high level of noise in the displacements for some segments appears in the results, which we can be improved in our future work by employing a corner cube reflector of a better quality. In addition, we could even improve the Leica MS50 results by taking angular tilt axis errors into account and by avoiding the very minor shaking of the instrument.

Conclusion

A robust and consistent procedure was proposed to perform an accurate displacement and vibration analysis of a footbridge structure using an IATS (here, Leica MS50). The Leica MS50 benefits accurate distance measurements to the object in addition to the captured video streams with a practical sampling frequency of 10 Hz using an embedded on-axis telescope camera. The experiments were carried out for two case studies under a controlled excitation in the laboratory environment and an uncontrolled excitation of a footbridge structure. In a first case study, the results were validated by means of two highly accurate reference measurement systems, namely, the portable shaker vibration calibrator 9210D (with a sampling frequency of 200 Hz) and the Leica AT960-LR LT (with a sampling frequency of 200 Hz). In the second case study, the validation was performed based on the known natural frequencies of the footbridge structure calculated from the FEM analysis. In addition, the LT measurement was also used for a validation.

To extract target centroid from video streams of the Leica MS50, the feasibility of an optimal passive target pattern including four intersected lines and its accurate and reliable detection approach, proposed in Omidalizarandi et al.,[26] were investigated at different

epochs of time as a preliminary step. Subsequently, the vertical angular conversion factor was calibrated in the laboratory environment to convert derived displacements from the pixel unit to the metric unit.

To estimate amplitudes and frequencies for all three sensors with high accuracy, vibration measurements either in the length unit or the acceleration unit were input to the Fourier series as a linear regression model comprising a sum of sinusoids and additive random deviations. Furthermore, the coloured measurement noise was decorrelated through a covariance-stationary AR process of an order 25, assuming the white noise components to independently follow a central and scaled t-distribution with an unknown scale factor and unknown degree of freedom. At the end, model parameters were estimated by means of the GEM algorithm as described in Alkhatib et al.[30] The unknown frequencies were initiated by means of notable maximum amplitudes within the DFT of the data to perform adjustment of the combined observation model for a footbridge application.

The results indicate that the estimated frequencies and amplitudes from the Leica MS50 measurements were very close to those resulting from the two highly accurate reference sensors (LT and PSVC) in the laboratory environment and to those resulting from the FEM analysis and the LT for the real application of the footbridge structure. It was shown that the DFT results and our proposed approach achieved approximately similar results when estimating the frequencies. However, the results for the amplitudes varied from minor to significant changes depending on the coloured noise behaviour of the measurements. To show the strength of the proposed approach, the estimated results were compared in two cases of the AR model in the order of 25: with an unknown degree of freedom and without the AR process considering a constant degree of freedom of 120, which is close approximation of a normal distribution.

The estimated degrees of freedom of the t-distribution, in the case of considering the AR model order of 25, reveals that the LT and the PSVC measurements are a rather close approximation of a normal distribution, while, by contrast, the estimated degrees of freedom regarding the Leica MS50 measurements with substantial tails show a large number of outliers in the measurement noise of that sensor. Moreover, the results show that the highest white noise test acceptance rate (100%) was obtained for the PSVC measurements. The LT measurements also produced relatively high acceptance rates in comparison to the PSVC data. However, the acceptance rates regarding the Leica MS50 measurements fluctuate between 25% and 75%, so that the adjusted coloured noise model is clearly inadequate for a number of analysed segments. Furthermore, the image motion error for some video frames derived from

a weak PC performance or delay in the data transmission procedure from the Leica MS50 to the PC has a significant influence on the estimated frequencies and amplitudes and shows a higher coloured noise level compared to the good quality video frames captured. In summary, the results show the feasibility of Leica MS50 for an accurate displacement and vibration analysis of the footbridge structure for frequencies less than 5 Hz (in view of the Nyquist sampling theorem).

In our future work, we will measure the minor shaking of the Leica MS50 throughout the measurements by continuously reading the tilting axis error of the Leica MS50 using the GeoCOM interface. It might be beneficial to increase the measurement time to obtain more redundant data and to be able to increase AR model order to improve the coloured noise models and the resulting performance of the white noise test. In addition, the proposed approach can be extended to the time-dependant AR model to characterize the coloured noise behaviour of the measurements over time. Furthermore, the time synchronization for the measurements of the sensors can be performed to have a more realistic comparison of the results at a certain point of time. Moreover, we will improve the LT measurements for a footbridge structure by employing a corner cube reflector of a better quality. The internal calibration of the error sources of the Leica MS50 measurements might improve the results. The possibility of performing experiment with the Leica MS60 with a maximum sampling frequency of 20 Hz could also improve the results. Finally, global optimization can be applied for more accurate and reliable results to estimate model parameters with unknown frequencies.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

collaborative project 'Spatio-temporal monitoring of bridge structures using low cost sensors' with ALLSAT GmbH, which was supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) and the Central Innovation Programme for SMEs (Grant ZIM Kooperationsprojekt, ZF4081803DB6).

ORCID iD

Mohammad Omidalizarandi https://orcid.org/0000-0002-9897-4473
Boris Kargoll https://orcid.org/0000-0003-3496-6391
Jens-André Paffenholz https://orcid.org/0000-0003-1222-5568

References

1. Alvandi A and Cremona C. Assessment of vibration-based damage identification techniques. *J Sound Vib* 2006; 292: 179–202.

2. Ehrhart M and Lienhart W. Monitoring of civil engineering structures using a state-of-the-art image assisted total station. *J Appl Geodes* 2015; 9: 174–182.

3. Lienhart W, Ehrhart M and Grick M. High frequent total station measurements for the monitoring of bridge vibrations. *J Appl Geodes* 2017; 11: 1–8.

4. Leach M and Hyzak M. GPS structural monitoring as applied to a cable-stayed suspension bridge. In: *Proceedings of the International Federation of Surveyors (FIG) XX congress, Commission 6*, Melbourne, VIC, Australia, 1994. [CD-ROM].

5. Psimoulis PA and Stiros SC. Measuring deflections of a short-span railway bridge using a robotic total station. *J Bridge Eng* 2013; 18: 182–185.

6. Pytharouli SI and Stiros SC. Spectral analysis of unevenly spaced or discontinuous data using the 'norm-period' code. *Comput Struct* 2008; 86: 190–196.

7. Roberts GW, Meng X and Dodson AH. Integrating a global positioning system and accelerometers to monitor the deflection of bridges. *J Surv Eng: ASCE* 2004; 130: 65–72.

8. Neitzel F, Niemeier W, Weisbrich S, et al. Investigation of low-cost accelerometer, terrestrial laser scanner and ground-based radar interferometer for vibration monitoring of bridges. In: *Proceedings of the 6th European workshop on structural health monitoring*, Dresden, 3–6 July 2012, pp.542–551. Deutsche Gesellschaft für Zerstörungsfreie Prüfung (DGZfP).

9. Vosselman G and Maas HG. *Airborne and terrestrial laser scanning*. Boca Raton, FL: CRC Press, 2010.

10. Omidalizarandi M and Neumann I. Comparison of target-and mutual information based calibration of terrestrial laser scanner and digital camera for deformation monitoring. *ISPRS Arch* 2015; 40: 559–564.

11. Omidalizarandi M, Paffenholz J-A, Stenz U, et al. Highly accurate extrinsic calibration of terrestrial laser scanner and digital camera for structural monitoring applications. In: *Proceedings of the 3rd joint international symposium on deformation monitoring (JISDM)*, Vienna, 30 March–1 April 2016, p.8. [CD-ROM].

12. Reiterer A, Lehmann M, Miljanovic M, et al. A 3D optical deformation measurement system supported by knowledge-based and learning techniques. *J Appl Geodes* 2009; 3: 1–13.

13. Bürki B, Guillaume S, Sorber P, et al. DAEDALUS: a versatile usable digital clip-on measuring system for total stations. In: *Proceedings of the international conference on indoor positioning and indoor navigation (IPIN)*, Zurich, 15–17 September 2010, pp.1–10. New York: IEEE.

14. Wagner A, Wasmeier P, Reith C, et al. Bridge monitoring by means of video-tacheometer – a case study. *Allgem Vermessungs-Nachr* 2013; 120: 283–292.

15. Wagner A, Huber B, Wiedemann W, et al. Long-range geo-monitoring using image assisted total stations. *J Appl Geodes* 2014; 8: 223–234.

16. Ehrhart M and Lienhart W. Image-based dynamic deformation monitoring of civil engineering structures from long ranges. *Proc SPIE* 2015; 9405: 94050J.

17. Guillaume S, Clerc J, Leyder C, et al. Contribution of the image-assisted theodolite system QDaedalus to geodetic static and dynamic deformation monitoring. In: *Proceedings of the 3rd joint international symposium on deformation monitoring (JISDM)*, Vienna, 30 March–1 April 2016. [CD-ROM].

18. Wagner A. A new approach for geo-monitoring using modern total stations and RGB + D images. *Measurement* 2016; 82: 64–74.

19. Wagner A. *New geodetic monitoring approaches using image assisted total stations*. PhD Thesis, Technische Universität München, München, 2017.

20. Leica Geosystems. *Leica Nova MS50 Geocom reference manual* (version 5.50). St. Gallen: Leica Geosystems AG, 2014.

21. Gruen A. Least squares matching: a fundamental measurement algorithm. In: Atkinson KB (ed.) *Close range photogrammetry and machine vision*. Dunbeath: Whittles Publishing, 1996, pp.217–255.

22. Akca D. Full automatic registration of laser scanner point clouds. In: *Proceedings of the optical 3-D measurement techniques VI*, Zurich, 22–25 September 2003, pp.330–337. Wichmann.

23. Gruen A and Akca D. Least squares 3D surface and curve matching. *ISPRS J Photogramm* 2005; 59: 151–174.

24. Zhou Y, Wagner A, Wunderlich T, et al. Close range angles deformation monitoring by telescope camera of total station MS50 using automatic detection of coded targets. In: *Proceedings of the 2nd international workshop on civil engineering and architecture*, Istanbul, 12–13 September 2015, pp.153–162.

25. Reiterer A and Wagner A. System consideration of an image assisted total station – evaluation and assessment. *Allgem Vermessungs-Nachr* 2012; 119: 83–94.

26. Omidalizarandi M, Paffenholz J-A and Neumann I. Automatic and accurate passive target centroid detection for applications in engineering geodesy. *Surv Rev.* 2018; 1–16.

27. Walser B. *Development and calibration of an image assisted total station*. PhD Thesis, ETH Zürich, Zürich, 2004.

28. Ehrhart M and Lienhart W. Accurate measurements with image-assisted total stations and their prerequisites. *J Surv Eng: ASCE* 2016; 143: 04016024.

29. Zhou Y, Wagner A, Wunderlich T, et al. Calibration method for IATS and application in multi-target monitoring using coded targets. *J Appl Geodes* 2017; 11: 99–106.

30. Alkhatib H, Kargoll B and Paffenholz J-A. Robust multivariate time series analysis in nonlinear models with autoregressive and t-distributed errors. In: *Proceedings of the international work-conference on time series analysis (ITISE)*, Granada, 18–20 September 2017. Godel Impresiones Digitales S.L.

31. Leica Geosystems. *Leica Nova MS50 white paper*. St. Gallen: Leica Geosystems AG, 2013.

32. Leica Geosystems. PCMM system specifications: Leica absolute tracker and Leica T-products, 2010, http://www.hsands.com/productpdfs/leica.pdf

33. The Modal Shop. Model 9210D: low frequency portable vibration calibrator (datasheet), 2015, http://www.modalshop.com/filelibrary/9210D-Low-Frequency-Portable-Vibration-Calibrator(DS-0140).pdf

34. Grompone von Gioi R, Jakubowicz J, Morel JM, et al. LSD: a line segment detector. *Image Process On Line* 2012; 2: 35–55.

35. Kaehler A and Bradski G. *Learning OpenCV 3: computer vision in C + + with the OpenCV library*. Sebastopol, CA: O'Reilly Media, Inc., 2016.

36. Canny J. A computational approach to edge detection. *IEEE T Pattern Anal* 1986; 8: 679–698.

37. PhysicsNet, http://physicsnet.co.uk (accessed 1 September 2017).

38. Andrews LC and Phillips RL. *Mathematical techniques for engineers and scientists*. Bellingham, WA: SPIE Press, 2003.

39. Kargoll B, Omidalizarandi M, Loth I, et al. An iteratively reweighted least-squares approach to adaptive robust adjustment of parameters in linear regression models with autoregressive and t-distributed deviations. *J Geodesy* 2018; 92: 271–297.

40. Abramowitz M and Stegun IA. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology, 1964.

41. Koch KR. Expectation maximization algorithm and its minimal detectable outliers. *Stud Geophys Geod* 2017; 61: 1–18.