IDEA

Innovations Deserving
Exploratory Analysis Programs

NCHRP IDEA Program

A Novel Vision Sensor for Remote Measurement of Bridge Displacement

Final Report for
NCHRP IDEA Project 189

Prepared by:
Maria Q. Feng
Columbia University

January 2019

TRB TRANSPORTATION RESEARCH BOARD
The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Innovations Deserving Exploratory Analysis (IDEA) Programs Managed by the Transportation Research Board

This IDEA project was funded by the NCHRP IDEA Program.

The TRB currently manages the following three IDEA programs:

- The NCHRP IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded by American Association of State Highway and Transportation Officials (AASHTO) as part of the National Cooperative Highway Research Program (NCHRP).
- The Safety IDEA Program currently focuses on innovative approaches for improving railroad safety or performance. The program is currently funded by the Federal Railroad Administration (FRA). The program was previously jointly funded by the Federal Motor Carrier Safety Administration (FMCSA) and the FRA.
- The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).

Management of the three IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies.

For information on the IDEA programs, check the IDEA website (www.trb.org/idea). For questions, contact the IDEA programs office by telephone at (202) 334-3310.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

A NOVEL VISION SENSOR FOR REMOTE MEASUREMENT OF BRIDGE DISPLACEMENT

NCHRP-IDEA Project 189

Final Report

By

Maria Q. Feng, Renwick Professor
Department of Civil Engineering and Engineering Mechanics
and her team at
Sensing, Monitoring, and Robotics Technology (SMaRT) Laboratory
Columbia University

January 19, 2019

SUMMARY

Bridge dynamic responses can reveal significant information about the structural health condition. This project develops a camera-based computer vision sensor system for accurate remote measurement of multipoint bridge displacements in outdoor environments. The vision sensor tracks the movement of natural targets on the structure from a convenient remote location without accessing the bridge to install sensors, and thus the cost of the equipment and operation is much lower than conventional sensors. In addition, the measurement points can be easily altered in post processing of the video. However, outdoor conditions such as changes in illumination and background, heat haze-induced image distortions, and camera vibration can cause significant measurement errors. To overcome these challenges, this project develops innovative algorithms and methods including (1) a gradient-based template matching algorithm for robust tracking of "natural" markers (such as rivets on a bridge), integrated with a subpixel technique for enhancing image resolution required for multipoint measurement; (2) a heat haze filtering technique based on detection and statistical characterization of heat haze-induced image distortions; and (3) a practical vibration cancellation method based on simultaneous measurement of displacements of the target structure and a stationary point. In addition, this project proposes a practical calibration method to convert image pixel displacements into physical displacements.

This project is executed in two stages. Stage 1 focuses on developing the algorithms, techniques and methods to systematically address all sources of environmental noise that deteriorate image quality and measurement accuracy. Extensive laboratory tests are carried out to validate the effectiveness of the developed algorithms and techniques using simulated environmental noise including low-lighting, shadowing, heat haze, and change in illuminating light and background conditions. In Stage 2, the algorithms developed in Stage 1 are integrated into a software package and field performance evaluation tests are carried in three bridges including two long-span steel bridges, the Manhattan Bridge and the Williamsburg Bridge, and a short-span concrete bridge, the Jamboree Bridge. The remote, real-time and multipoint measurement capabilities of the vision sensor system developed in this study are further validated in presence of various sources of field environmental noise including heat haze and camera vibration.

In addition, this project demonstrates, through a laboratory experiment, the use of the vision sensor system for structural dynamic tests, modal analysis, and damage detection, by simultaneously measuring a dense array of displacements at extremely low cost. In the future, the system can be further developed for permanent installation at bridge sites to enable long-term continuous monitoring of structural integrity and safety.

TABLE OF CONTENTS

1. INTRODUCTION

The average age of the nation's bridges is approaching 50 years and 9.1% of the bridges are rated structurally deficient (ASCE, 2017). According to AASHTO (2008), a bridge structure is categorized as structurally deficient when "significant load-carrying elements are found to be in poor condition due to deterioration". It has become increasingly important to properly inspect, monitor and maintain these aging and deteriorating bridges. The Manual for Bridge Evaluation by AASHTO (2010) recommends load testing to rate the bridge's load-carrying capacity by measuring bridge displacements under given loads.

Such practice, however, is highly expensive to operate, mainly due to cumbersome, time-consuming, and expensive installation of sensors and their data acquisition systems. Sensors currently available for measuring bridge displacements can be classified as contact type (e.g., LVDT) and non-contact type (e.g., laser vibrometer) sensors. The contact sensor requires the access to the bridge to install the sensor and physically connect it to a stationary reference point, which is often difficult or even impossible to locate, particularly for bridges crossing over a river, a road or deep/rugged terrain. Moreover, these sensors need to be cabled to external devices for data acquisition (DAQ) and power-supply, making the installation even more cumbersome, time-consuming and costly. The GPS sensors are easier to install because it does not require a stationary reference point, but the measurement accuracy is limited, usually with accuracy in the range of 5mm to 10mm.

The non-contact laser vibrometer is accurate, but it is costly and has limited measurement distance, because a longer distance measurement would require the use of a higher-intensity laser beam that could endanger human health (Nassif et al. 2005; Ribeiro et al. 2014). The interferometric radar provides a comparable performance, but requires reflecting surfaces to be mounted on the measurement points of the structure, which significantly limits its applications (Gentile and Bernardini 2009; Su et al. 2012). In addition, the costs of these systems are high.

The rapid advances in cameras and computer vision techniques enable the camera-based vision sensor a promising alternative to conventional sensors for structural displacement measurement and health monitoring. Simply place a camera at a convenient location and record the video of a target structure. The structural displacement time histories can be extracted from the frames of the video images by tracking the movement of selected target points on the structure. The emerging vision sensor technology offers significant advantages over the contact-type and other noncontact-type (e.g., GPS, laser vibrometer) displacement sensors, as summarized below:

1) In contrast to the conventional contact-type sensor (such as an LVDT) that requires time-consuming and costly installation of the sensor on the structure with physical connections to not only a stationary reference point but also DAQ and power supply, the vision sensor does not require physical access to the structures, as the camera can be set up at a remote location. This represents a significant time and cost saving. For bridge monitoring, for example, no traffic control is required.
2) Compared with the GPS, which still requires installation on the structure (but not the stationary reference point), the vision sensor is far more accurate and less expensive. Depending on the cost, the GPS measurement error is typically in the range of 5 mm to 10 mm, more than an order of magnitude larger than that of the vision sensor.
3) Compared with the non-contact laser vibrometer, which needs to be placed relatively close to the measurement target due to the low laser power for safety concerns, the vision sensor can be placed hundreds of meters away (when using a zoom lens) and still achieve satisfactory measurement accuracy.
4) In contrast to these conventional sensors, all of which are point-wise sensors, the vision sensor can be termed as a noncontact distributed sensing technique as it can simultaneously tracking multiple points from a long distance. More importantly, one can easily alter the measurement points after the video images are taken.

The PI and her team started to develop this emerging technology in 2004 for measuring bridge structural displacements for the first time. Initially, they developed a low-cost vision-based displacement sensor system that required a target panel with a predesigned high-contrast pattern to be attached to the bridge for accurate tracking (Lee and Shinozuka 2006; Fukuda et al. 2010). To eliminate the requirement for bridge access to install the target panel, they improved the target tracking algorithm and succeeded in tracking a "natural" target on the bridge surface for displacement measurement (Fukuda et al. 2013, Feng et al. 2015). But this work was limited to measuring a single point displacement with one camera. Ribeiro et al. (2014) measured the dynamic displacements of a railway bridge based on video images. The achieved measurement accuracy was 0.1mm for a measurement distances (from the camera to the target) up to 15m, but the accuracy deteriorated as the measurement distance increased. It is noted that this study measured a single point and required the installation of a pre-designed high-contrast target panel for easy tracking. Also, depending on the field conditions, it is often impossible to set up the camera within several meters from the bridge's measurement points. Busca et al.(2014) attempted multiple-point measurement of dynamic bridge displacements by using different types of targets at the measurement points. However, without a high-contrast target marker, they found that it is highly difficult to accurately extract displacements when the images are taken in the field where the shading, lighting, and background conditions change.

For bridge monitoring applications, it is particularly challenging to accurately measure multi-point bridge displacements using one camera placed from distance and without artificial targets, due to the limited image resolution for each measurement point and more importantly, various optical noise in the field such as insufficient and changing light conditions, heat haze, and camera vibration caused by wind and traffic. Building upon the prior work, this project aims at developing a computer vision system for *accurate, long-distance, outdoor* measurement of multipoint bridge displacements in difficult outdoor environments involving changes in illumination/background, heat haze-induced image distortions, and wind/traffic-induced camera vibration.

2. VISION SENSOR SYSTEM AND PROJECT OVERVIEW

The proposed computer vision system for measuring bridge displacements is illustrated in Fig. 1. The system simply consists of a laptop or tablet computer installed with measurement software and connected to one or multiple video cameras (e.g., CCD/CMOS image sensor with zoom lens, camcorder, DSLR camera, security camera or even a low-cost smartphone camera), which can measure structural displacements at multiple locations simultaneously. The video images of natural targets on a structure, such rivets and edges, captured by the camera are digitized and streamed into the computer. No artificial high-contrast markers need to be installed on the target bridge. Instead, the displacements of the bridge are measured by tracking the movement of existing "natural" markers such as rivets or edges through template matching techniques. As the camera records the video of the target bridge from a distance, the displacements of these "natural" markers, i.e., the bridge displacements, can be extracted in real time from the images by the built-in software. The images can also be saved for post-processing, in which the measurement points can be altered, providing flexibility of changing measurement points, which cannot be done with conventional sensors installed at fixed locations.

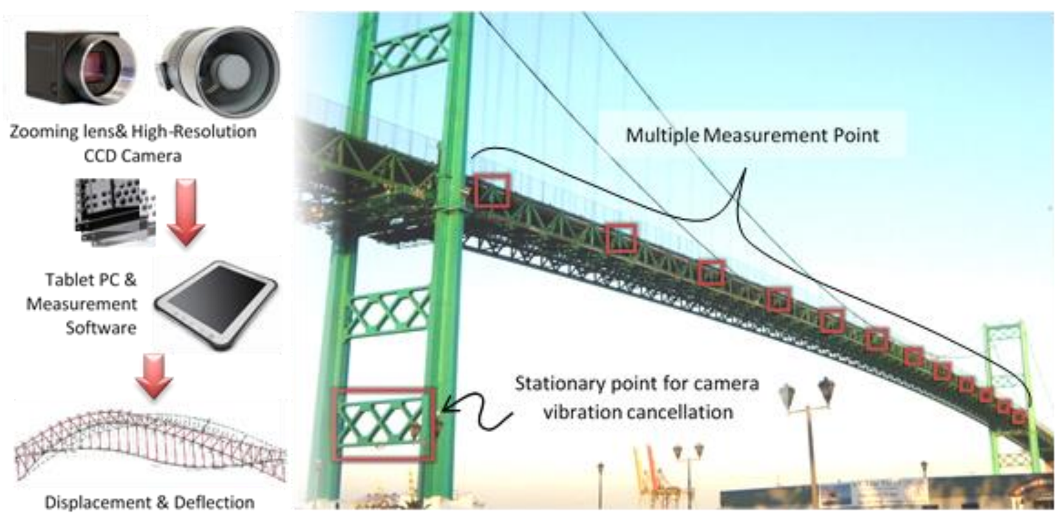


Figure 1. Vision sensors for remote multipoint measurement of bridge displacements

The computer vision sensor system is easy to set up, inexpensive to operate, and does not require access to the bridge when natural targets are used. However, implementation of the computer vision sensor system for accurate measurements in the field has been challenging.

This IDEA project systematically addresses vision measurement errors caused by the various sources of outdoor environmental noise through the innovations made in three areas. (1) A novel target tracking algorithm, referred to as Orientation Code Matching (OCM) created by the project team, is further developed to enable robust tracking of natural targets on the bridges and simultaneous measurement of displacements at multiple points with one camera placed at a convenient remote location. It is found that the OCM scheme is more robust than conventional schemes using image intensity values for template matching which relies on the image quality, because OCM employs the gradient information in the form of orientation codes which is inherently invariant to variations in image intensity and thus more robust when irregularities are present., (2) An statistical filtering technique is developed to overcome image distortions caused by heat haze-induced variation in the light reflection index of air, which decreases the displacement errors for long-distance measurement. (3) A practical vibration cancellation technique is developed to minimize measurement errors caused by camera vibration, which, again, particularly benefit long-distance measurement in which the camera vibration is amplified by the zoom lens. In addition, this project also develops a practical calibration method to accurately convert pixel displacements to physical displacements. These innovations enable accurate simultaneous measurement of bridge displacements at multiple points using one camera from a long distance in outdoor environments without requiring the installation of artificial targets on the bridge, as demonstrated in extensive laboratory and field tests on two long-span steel bridges, the Manhattan Bridge and the Williamsburg Bridge, and a short-span concrete bridge, the Jamboree bridge.

3. VISION-BASED DISPLACEMENT MEASUREMENT TECHNIQUES

Most of the existing computer vision-based sensors suffer from one or more of the following limitations for practical applications: (1) The accuracy of the template matching techniques is largely dependent on the image quality, which is often difficult to guarantee in outdoor field environmental conditions such as illumination variation, partial target occlusion, partial shading, and background disturbance, etc.. (2) The adopted template matching techniques gives displacement in integer-pixel resolution since the minimal unit in a video image is one pixel. Although in many applications the pixel-level accuracy is adequate, it is often far from the required accuracy when simultaneously measuring multiple points of small structural vibrations with one camera. (3) In the field environment, it is challenging to determine the scaling/calibration factor that can accurately converts the image coordinates in the unit of pixels into physical coordinates.

This chapter addresses these practical limitations by developing a real-time image processing algorithm and software based on the robust OCM template matching scheme, a subpixel technique, and a practical calibration method. The software is evaluated in laboratory experiments involving a simply supported beam without and without structural damage. By accurately measuring the structural vibration at multiple points using one camera, the structural damage, together with its location, is successfully detected.

3.1 Robust Multipoint Displacement Measurement Software

This project developed a robust template matching algorithm based on the OCM scheme enhanced by a subpixel technique for multipoint displacement measurement. The algorithm was further developed into a software package for real-time displacement measurement. The algorithm is illustrated in Fig. 2. In the OCM scheme, orientation code representations of both the object and template images are constructed from the corresponding gray intensity images. In this way, each pixel represents an orientation code obtained by quantizing the orientation angle at the corresponding pixel position in the original images (Feng, D.M., 2016).

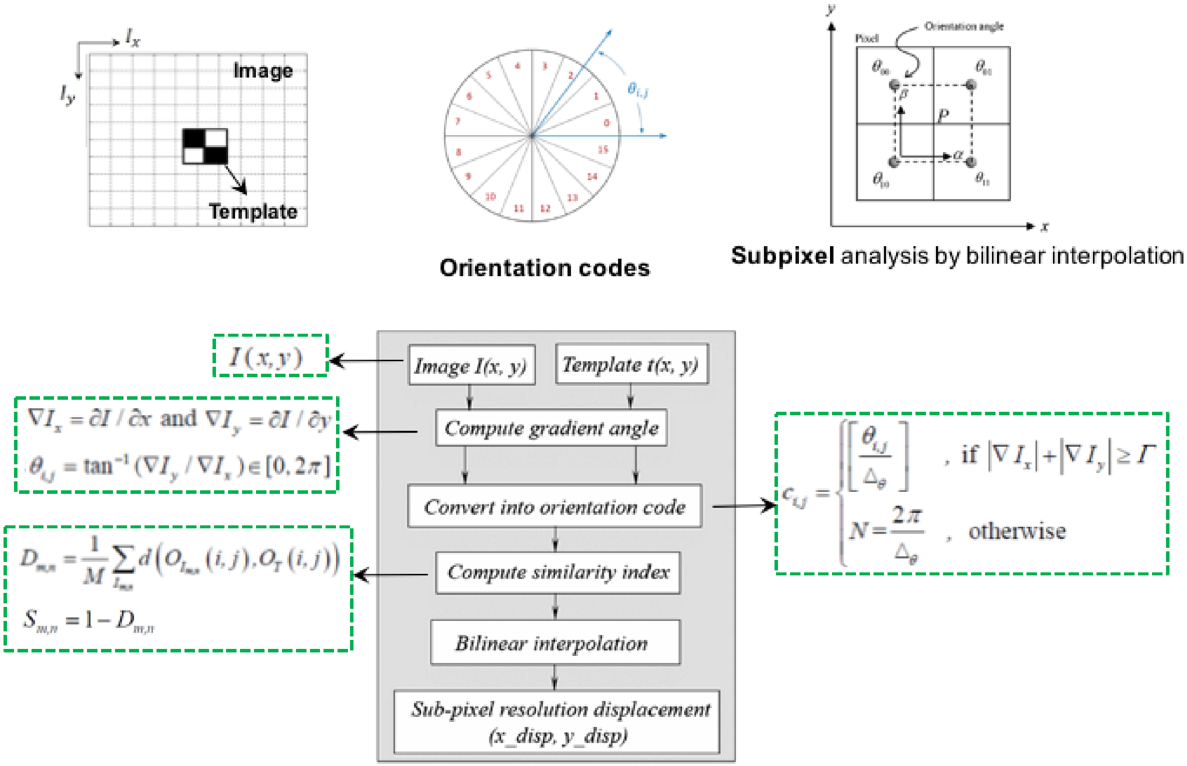


Figure 2. Subpixel-enhanced OCM algorithm

In the implementation, an initial area in the first image of a sequence of video frames captured is defined as a template. The template can be located in the successive images using the template matching techniques. Thus the displacement in pixels is obtained, which can be further transformed into physical displacement through a scaling factor. To improve the image resolution, subpixel registrations are incorporated into the template matching algorithm.

Pixel-level template matching may result in unacceptable measurement errors if the displacement to be measured has same order of magnitude as the scaling factor. This is particularly important for multipoint displacement measurement using a single camera. A subpixel technique can be adopted to make template matching fall at a fractional pixel location. This project further develops the OCM template matching algorithm by integrating the subpixel bilinear interpolation technique, as illustrated in Fig. 2.

The OCM template matching algorithm integrated with the subpixel technique is further developed into a real-time software package. The software tracks structural displacements at user-defined locations frame by frame. A user-friendly interface is built into the software package using Visual Studio 2010 with the C++ language, as shown in 3.

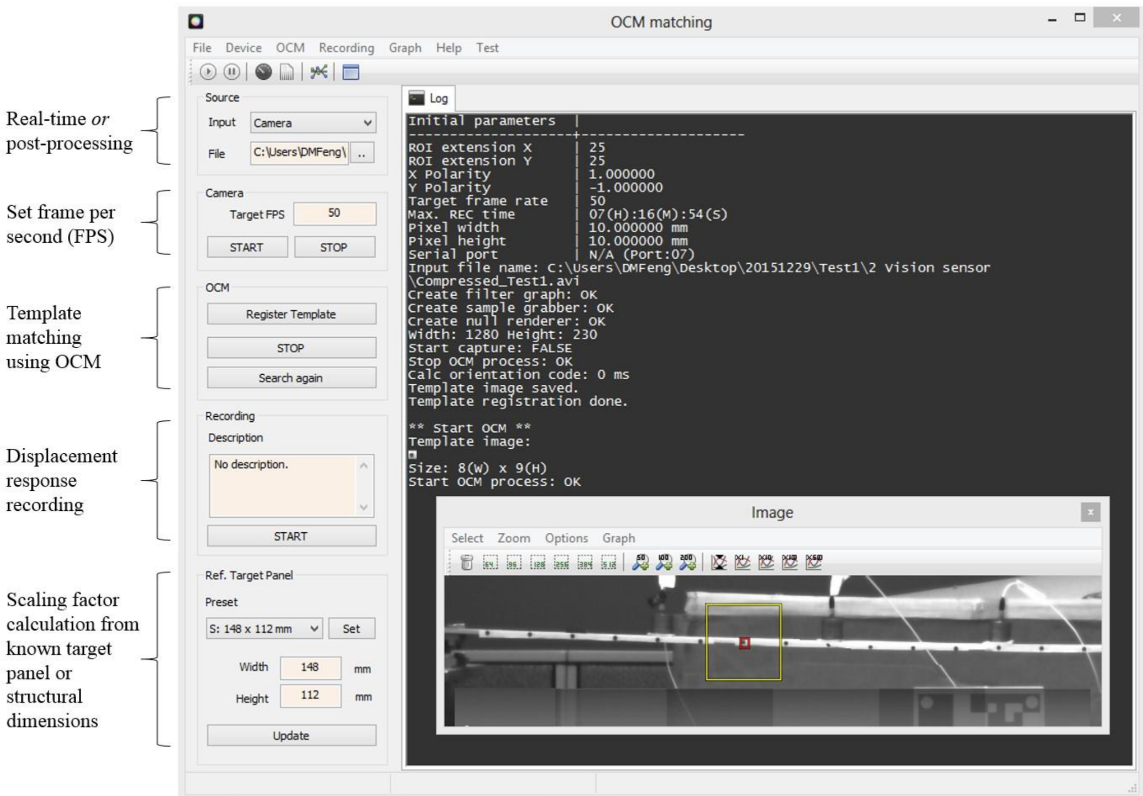


Figure 3. User interface of the OCM-based displacement measurement software

The procedure of real-time displacement measurement using the vision sensor system is described as follows:

1) *Vision sensor setup*. Fix the camera equipped with a zoom lens on a tripod and place it at a convenient location away from the structure. Through an USB 3.0 cable, the camera is connected to a laptop installed with the real-time displacement measurement software. It is noteworthy that setting up the vision sensor, including focusing the lens on the structural targets, takes only a few minutes.
2) *Calibration scaling factor determination*. In order to obtain structural displacements from

the captured video images, establishment of the relationship between the pixel coordinate and the physical coordinate is required. The scaling factor (e.g., with units of mm/pixel) can be obtained in two ways: (i) be estimated from the known physical dimension on the object surface and its corresponding image dimension in pixels; (ii) be estimated based on the intrinsic parameters of the camera as well as the extrinsic parameters between the camera and the object structure. In this study, the scaling factor is constructed using the first method, the advantage of which is discussed later. For example, if a 100mm long structural length is occupied with 500 pixels, the scaling factor would be 0.2mm/pixel. When structural dimension is not available, a reference target/marker panel with known dimension can be mounted on the structural surface.

3) *Single- or multiple-target/template registration*. Any texture on the structural surface can be registered as a tracking target, as long as it has pattern contrast compared with surrounding background, e.g., existing surface features such as bolt/rivet connection.
4) *Template matching for displacement extraction*. By clicking “START” on the “Recording” module in Fig. 3, video images captured by the camera are digitized into images with specified resolutions in 8 bit grey scales and streamed into the computer. Then OCM template matching algorithm together with the subpixel technique are employed to track the targets registered in the step 3). Thus the measured displacement history would be shown on the screen in real time and saved to the computer.

It is noted that in step 4), it would be highly time-consuming (thus making real-time measurement impossible) if the target is searched within the whole image of each video frame. To reduce computational time, the searching area could be confined to a predefined region of interest (ROI) near the template’s location in the previous image. It’s noted that the new ROI of a target must be able to cover its potential position on the next frame. Otherwise, mismatching will be introduced. It is also noted that geometrical distortion due to lens optics, especially when short focal length is used, would occur in the video images. In this case, a camera calibration process can be performed to reduce the effect of lens distortion. Herein, this study is mainly focusing on measuring small structural motions with relatively long focal-length lens from a remote distance. Thus the removal process of lens distortion is omitted.

The real-time displacement measurement avoids time-consuming and memory-intensive task of saving huge video files. However, tradeoffs among factors, such as measurement points, video resolution, maximum frame rate per second, template and ROI sizes, etc., are necessary. On the other hand, the developed software can also be used for post-processing the recorded video files. In this way, only a consumer-grade commercial video camera and a tripod are needed to take videos during laboratory experiments or outdoor field tests. This also enables the flexibility to extract structural displacements at any point from a single recording.

3.2 Practical Calibration Method

In order to obtain structural displacements from the captured video images, the establishment of the relationship between the pixel coordinate and the physical coordinate is required (e.g., with units of mm/pixel). When the camera optical axis is perpendicular to the object surface, all points on this surface have equal depth of fields, which means that these points can be equally scaled down into the image plane. In this case, only one identical scaling factor is needed. In general, the scaling factor can be obtained from one of the two methods, as expressed in Eq. **Error! Reference source not found.**: (1) SF_1 based on the known physical dimension on the object surface and its corresponding image dimension in pixels (i.e., d_{known} and I_{known}); (2) SF_2 based on the intrinsic parameters of the camera as well as the extrinsic parameters between the camera and the object structure (i.e., D, f and d_{pixel}). Particularly, when the camera optical axis is tilted about the normal directions of the object surface by an angle θ, as shown in Fig. 4, the scaling factor can be estimated by SF_3.

$$
\begin{aligned}
SF_1 &= \frac{d_{known}}{I_{known}} \\
SF_2 &= \frac{D}{f} d_{pixel} \\
SF_3 &= \frac{D}{f\cos^2\theta} d_{pixel}
\end{aligned}
\tag{1}
$$

where d_{known} is the known physical length on the object surface, I_{known} are the corresponding pixel length at the image plane, d_{pixel} is per pixel length (e.g., in μm/pixel), D is the distance between the camera and the object, and f is the focal length.

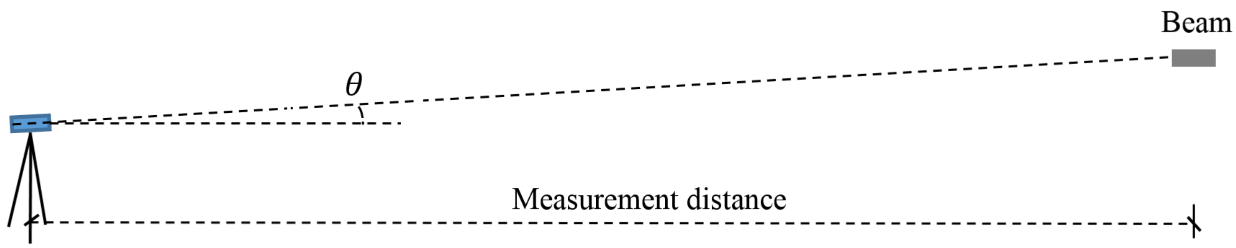


Figure 4. Case of non-perpendicular camera optical lens axis

For the scaling factor SF_3, estimation errors would arise from the uncertainties in the tilt angle estimation, camera distance measurement and focal length readings from the adjustable-focal-length lens. In such cases, the fixed zoom lens and angle measurement system can be used to minimize the error. Previous theoretical study found that for a fixed camera setup, the measurement error from scaling factor SF_1 would further decrease when the tracking target gets closer to the known structural dimension. Especially, the error is minimized when the tracking target is located within the region of known dimension.

3.3 Laboratory Evaluation Tests

In order to evaluate the efficacy of the vision sensor techniques and the displacement measurement software, a set of experiments are conducted using a frame structural model and a simply supported beam model in the SMaRT Laboratory at Columbia University. This tests further demonstrates the cost-effectiveness of the simultaneously measured structural displacements at multiple points for identification of structural modal parameters and detection of structural damage,

3.3.1 Evaluation of Robustness

In realistic field environments, various ill conditions such as illumination fluctuation, partial target occlusion, or background disturbance are often encountered, making it difficult to track the movement of a target, particularly a natural target. To investigate the robustness of the vision sensor under such environmental conditions, four testing cases are conducted using a two-story frame structure model mounted on a shaking table, as shown in Figure 5. The vision sensor system consists lap-top computer equipped with the developed image processing software and a video camera (Point Grey/FL3-U3-13Y3M-C) that has a CMOS-type sensor, with a maximum resolution of 1280×1024 pixel and maximum frame rate of 150 FPS. The optical lens has a focal length varying from 16-160mm with manual focus.

The vision sensor measures the frame displacements under sinusoidal excitations in four cases.

- Case 1: measurement by tracking an artificial target in normal light;
- Case 2: measurement by tracking a bolt connection in normal light;
- Case 3: measurement by tracking a bolt connection in dim light to simulate illumination fluctuation and background disturbance; and
- Case 4: measurement by tracking a bolt connection to simulate partial template occlusion.

The displacements measured by the OCM algorithm are compared with those by a conventional pixel intensity based digital image correlation method – the upsampled cross correlation (UCC) algorithm. UCC uses the cross correlation by means of Fourier transform in the spatial domain of the images.

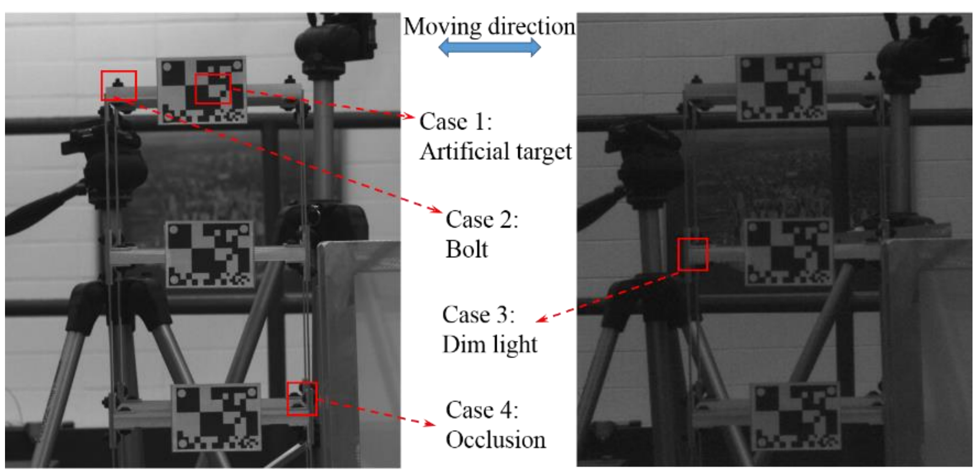


Figure 5. Test setup for evaluation of robustness against ill conditions

In the well illuminated condition, both OCM and UCC can accurately measure dynamic displacements. Figures 6 and 7 show the displacements of the top floor of the frame under the 1-Hz sinusoidal excitation, measured by remotely tracking the artificial target (Case 1) and the natural target – the bolt connection (Case 2) and processed with the OCM algorithm, in comparison with the UCC algorithm. Good agreements are observed and both UCC and OCM accurately measured the displacements, as shown in Figs. 6(a) and 7(a). In the contours of the correlation function by UCC, maximum values can be located at the true template matching positions, as shown in Figs. 6 (b) and 7 (b) for an example time instant of 2.5 sec

In the dim light condition, as in Case 3, OCM can successfully measure displacement but UCC completely fails to do so, as shown in Fig. 8 (a). The contour of the correlation function by UCC, as shown at an example time instant of 2.5 sec in Fig. 8 (b), presents a sequence of peaks, making it impossible to find the right template matching location.

When the target is partially occluded, as in Case 4 in which part of the template at the bottom right corner of the frame is occluded when the structure is moving to the right, OCM still can successfully measure the displacements, while UCC completely fails to track the template. The measured time histories by these two algorithms are in Fig. 9 (a). Figure 9(b) is the contour of the correlation function by UCC at an example time instant of 2.0 sec, in which multiple peaks appear, whose maximum is not located at the right matching point.

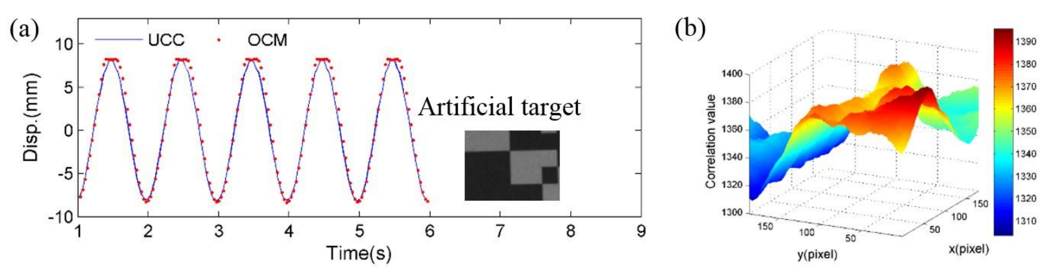


Figure 6. Case 1: (a) Displacements by OCM and UCC; (b) UCC cross correlation function contour

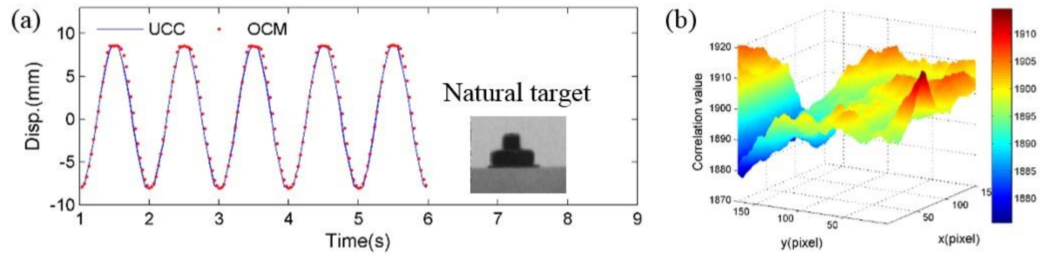


Figure 7. Case 2 (a) Displacements by OCM and UCC; (b) UCC cross correlation function contour

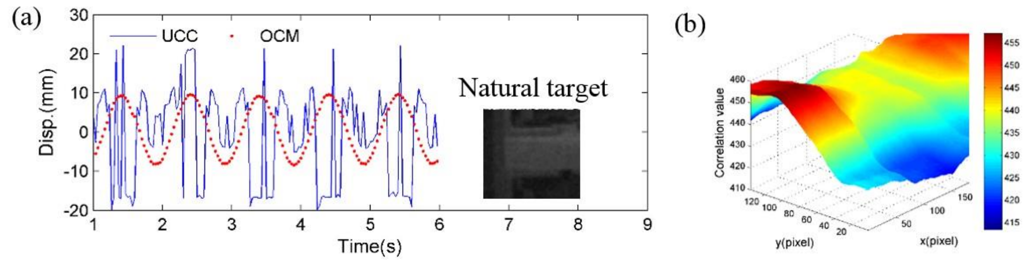


Figure 8. Case 3: (a) Displacements by OCM and UCC; (b) UCC cross correlation function contour

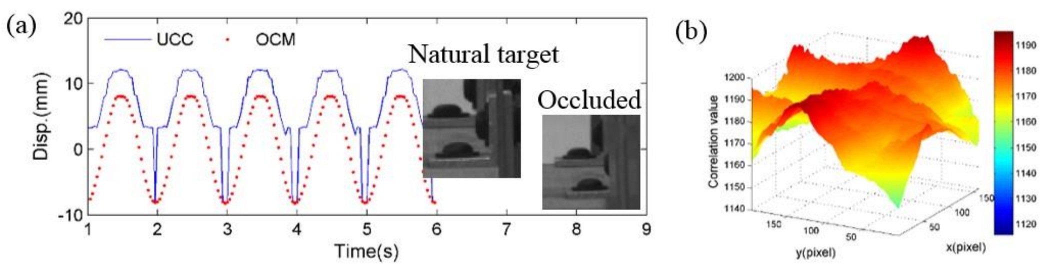


Figure 9. Case 4: (a) Displacements by OCM and UCC; (b) UCC cross correlation function contour

In summary, OCM is more robust in the ill conditions such as dim lights and partial occlusion of the images, which are common in outdoor field measurements. This is reasonable since UCC, the conventional template matching algorithm, utilizes the image intensity values for template matching which relies on the image quality, while OCM employs the gradient information in the form of orientation codes which is inherently invariant to variations in image intensity and thus more robust when irregularities are present.

3.3.2 Evaluation of Accuracy of Multipoint Measurement

In order to evaluate the accuracy of multipoint measurement of the OCM algorithm enhanced by the subpixel technique, a shaking table test is carried out on a three-story frame structural model. The frame is fixed on a shaking table (Model# APS113 by APS Dynamics Inc.) and subjected to white noise excitations. As shown in Fig. 10, four points on the frame model are measured by a single camera, which targets four black and white artificial targets (99 mm×75 mm) and four natural targets (bolt connections) on the frame. To evaluate the measurement accuracy, four high-fidelity laser displacement sensors (LDSs, Model#LK-G407 by KEYENCE)) are used as reference sensors. It is noted that the scaling factor for the vision sensor in the testing is 1.338mm/pixel, meaning the expected maximum error is 0.669mm from pixel-level template matching.

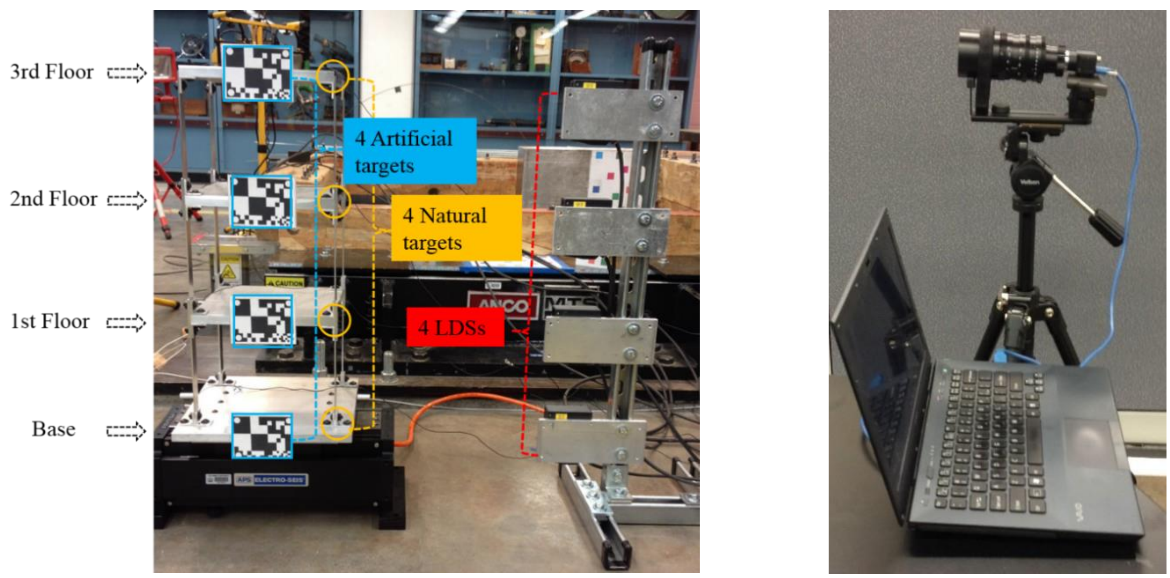


Figure 10. Laboratory test setup: (a) Shaking table test, (b) Vision sensor system

To quantify measurement accuracy, the normalized root mean squared error (NRMSE) is computed using the measured time histories by the vision sensor and the reference sensor.

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{y_{max} - y_{min}} \times 100\% \tag{2}$$

where n=number of measurement data; x_i and y_i=ith displacement data at time t_i, measured by the vision sensor and the reference sensor, respectively; and y_{max}=max(y_i), y_{min}=min(y_i).

Table 1 tabulates the measurement NRMSE errors for each floor of the frame structure computed from Eq. (2) based on the measured response displacements to random excitations from the shaking table over a measurement period of 5 minutes. While the displacement measurement by tracking the natural targets results in more errors than tracking the artificial targets, the maximum NRMSE error is merely 0.72%.

Table 1: Measurement errors: NRMSE (%)

Floor	OCM	
	Artificial target	Natural target
Base	0.35	0.72
1^{st}	0.24	0.56
2^{nd}	0.21	0.43
3^{rd}	0.14	0.35

3.3.3 Evaluation of the Calibration Method

When using a camera to monitor displacements at multiple points, the view angle change may affect the calibration accuracy. A simply supported beam model shown in Fig. 11 is used to evaluate the calibration the accuracy of the calibration methods proposed in Section 3.2. The same model is also used to test the multipoint measurement capability of the vision sensor and demonstrate its usefulness modal analysis and structural damage detection As shown in Fig. 11, 30 black dots, numbered from 2 through 31, are marked along the beam as targets for motion tracking. During the measurement, video images captured by the camera are digitized into 1280×240 pixel images in 8 bit grey scales and streamed into the computer with a sampling rate of 50 frames per second.

As references, the displacements are also measured by two reference sensors LDSs at point 9 and point 16, respectively. Besides, six accelerometers (Model#W352C67 by PCB PIEZOTRONICS Inc.) are installed to further compare the experimental modal analysis results. Both the LDSs and accelerometers adopt a sampling frequency of 50Hz.

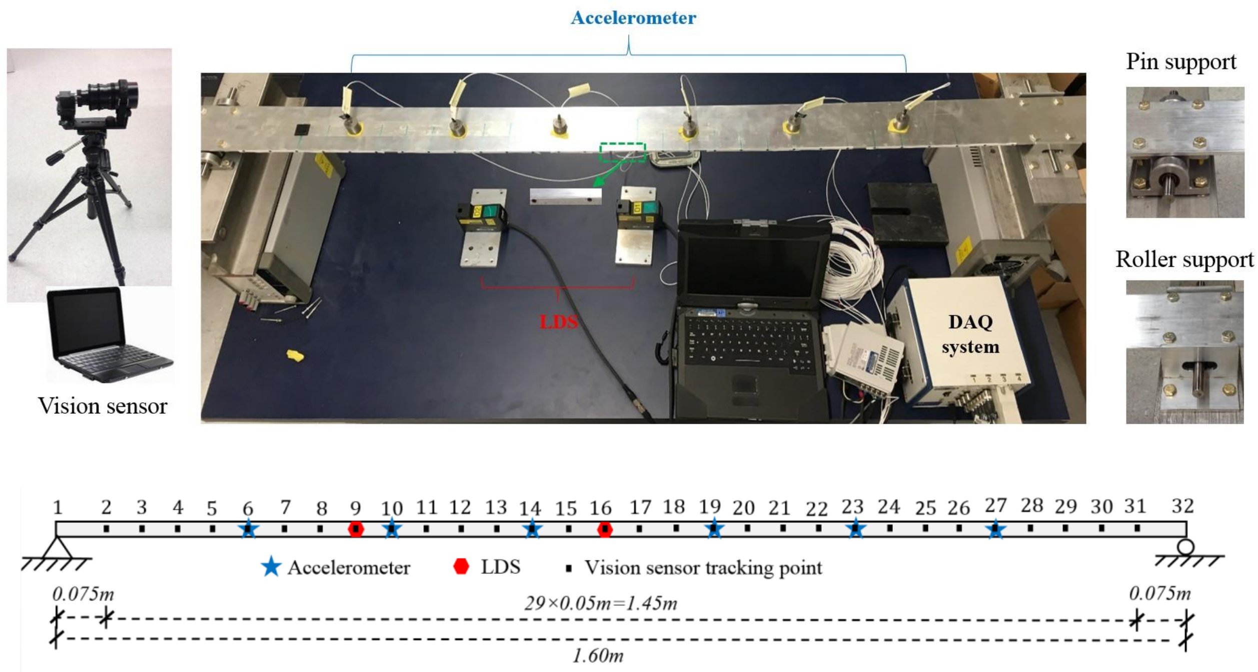
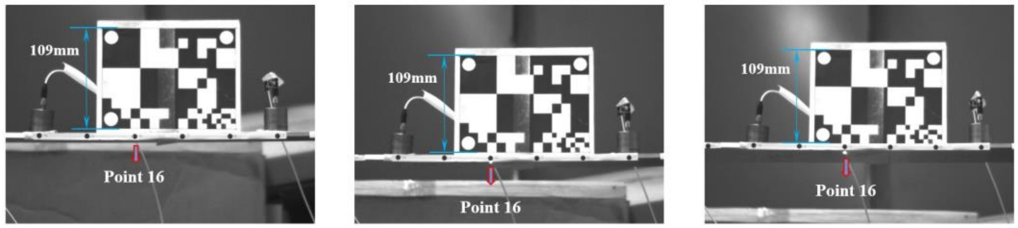


Figure 11. Test setup for evaluation of calibration and damage detection

The scaling factor is determined using SF_1 based on the known physical dimension on the object surface (e.g., the size of artificial target panels or the size of the nuts and rivets known from the design drawings) and the corresponding image dimension in pixels, as described in Eq. (1). It is noted that there exist uncertainties when picking the image dimension using a mouse. It is recommended to use the mean value from several repeated picking operations to average out some of the random errors. For the cases of non-perpendicular lens optical axis, scaling factors in the horizontal and vertical directions should be obtained respectively. Moreover, when a series of targets are tracked to measure displacements at multiple points along the structure, due to the projective distortion, different scaling factors should be determined by utilizing structural dimensions closer to or encompassing each of the targets. If such dimensions are not available, predesigned dimension panels can be attached as references, in which case, in order to minimize the estimation error of the scaling factors, measures should be taken to guarantee the reference panels are placed in the same plane of the object surface.



(a) (b) (c)

Figure 12. Tests of different camera tilt angles: (a) 3 , (b) 5 and (c) 9

To better illustrate this, the known vertical length (109mm) of a predesigned marker panel in Fig. 12 is used to estimate the vertical scaling factors for test cases of three camera tilt angles (i.e., 3 , 5 , 9). Scaling factors of 0.394mm/pixel, 0.407 mm/pixel and 0.426 mm/pixel are accordingly obtained.

Free vibration of the beam is generated by inducing an initial displacement at point 4. Displacement at point 16 is extracted by tracking the corresponding black dot and compared with that from LDS, as shown in Fig. 13. To evaluate the accuracy of the vision sensor, error analysis is carried out using the normalized root mean squared (NRMS) error of the displacements measured by the vision sensor and the LDS. The NRMS errors of the vision sensor corresponding to the three camera tilt angles in Fig. 13 are 1.0%, 1.2% and 1.4%, respectively. This demonstrates that the estimated scaling factors by utilizing known physical dimensions can guarantee satisfactory accuracy for cases of small camera tilt angles.

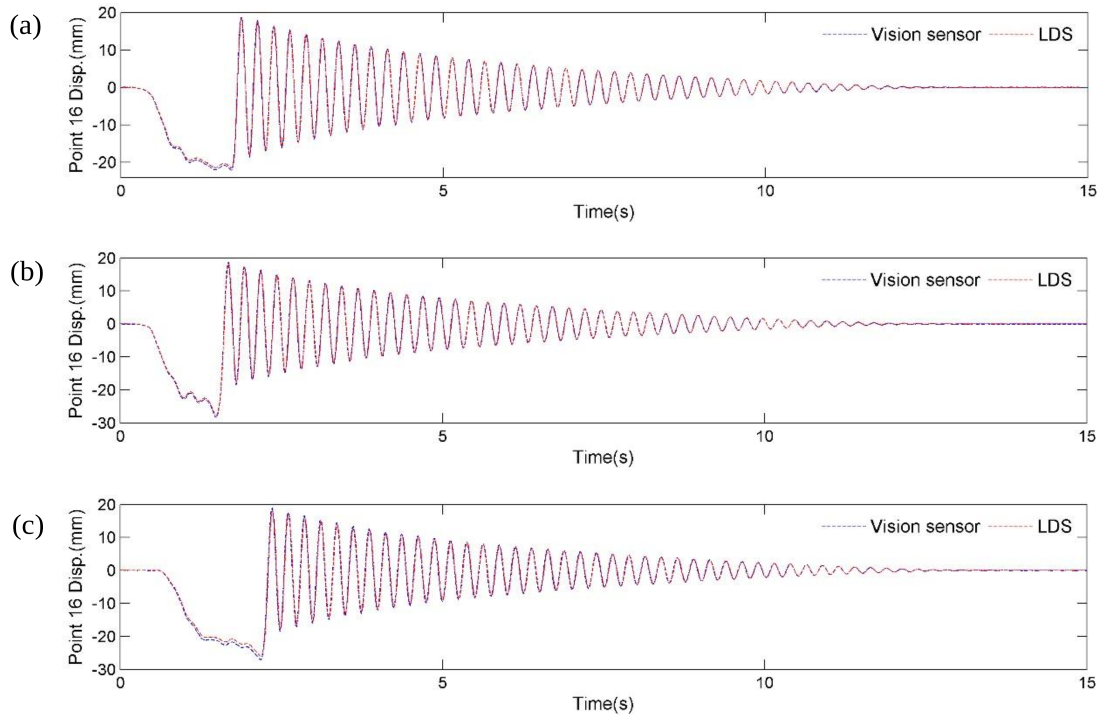


Figure 13. Comparison of displacement measurement at point 16 with optical lens angles: (a) 3°, (b) 5° and (c) 9°.

3.3.4 Multipoint Measurement for Modal Analysis Based on Free Vibration Test

To demonstrate the cost-effectiveness of the simultaneous measurement of multipoint displacements, experimental modal analysis is carried out using the vision sensor measurements. The first two theoretical natural frequencies of the model are computed as 4.52 Hz and 18.13 Hz.

For the following tests, the camera optical lens axis is placed to be perpendicular to the beam side surface. The scaling factor is 1.20mm/pixel for both horizontal and vertical directions.
By processing the digital video images on the computer with the developed image processing software, displacement time histories at points 2 through 31 are obtained. The displacements measured at point 9 and 16 by vision sensor are plotted and compared with those by LDSs in Fig.15, with the NRMS errors of 1.7% and 1.1%, respectively.

Figure 16 compares the extracted natural frequencies by the vision sensor, LDS and accelerometers, respectively. To better pick up the peaks at higher frequencies, a log scale is adopted. However, compared to the results in Figure (c), only the first natural frequency is obviously identified from displacement measurements by the vision sensor and the LDS, while the second natural frequency at 15.82Hz is difficult to identify. It is reasonable since accelerometer is inherently sensitive to high-order vibrations, while displacement sensor is essentially sensitive to low-frequency high-amplitude ones. In this case of the free vibration test, higher-order displacement components are very weakly excited. One solution is employing a higher-resolution camera sensor to increase the measurement resolution, i.e., obtaining a much smaller scaling factor to capture smaller-amplitude vibration components.

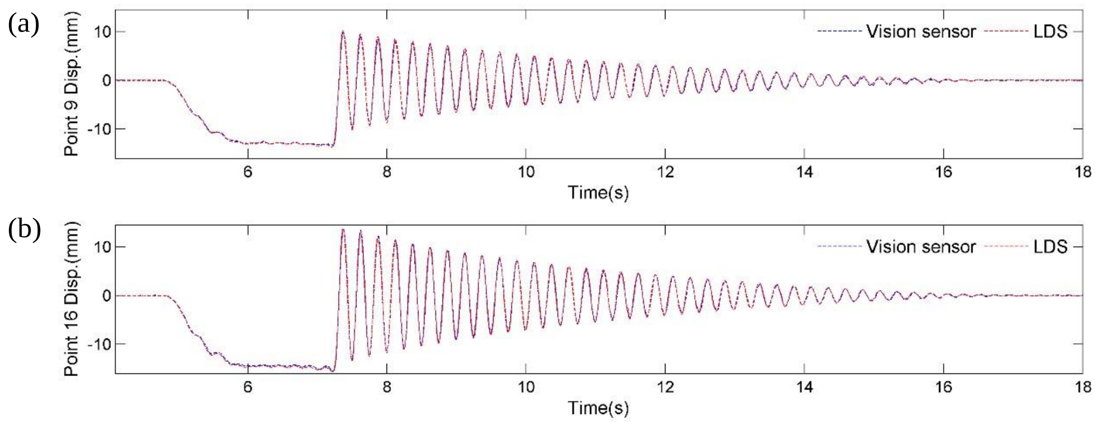
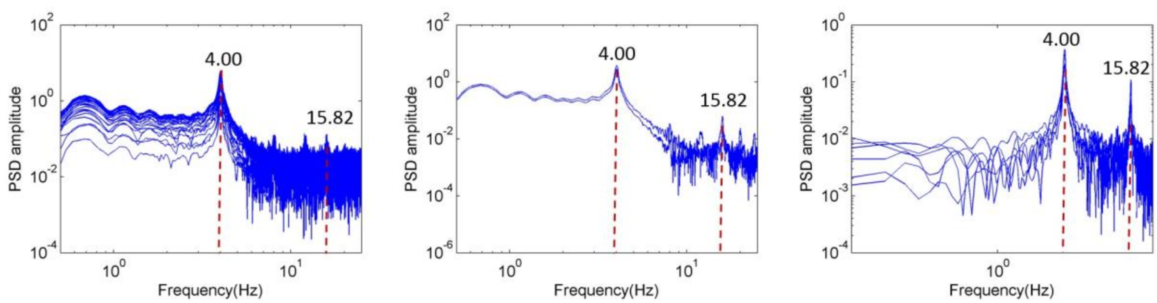


Figure 15. Comparison of displacement measurement: (a) at point 9, and (b) at point 16



(a) (b) (c)

Figure 16. Frequency results from: (a) displacement measurements at point 2 through 31 by vision sensor, (b) displacement measurements at points 9 and 16 by LDSs, and (c) accelerations at 6 points by accelerometers.

3.3.5 Multipoint Measurement for Modal Analysis Based on Hammer Impact Test

Herein, hammer impact is used to excite higher structural vibrational components. The time histories at 30 points along the intact beam for hammer hit at point 4 of the intact beam are shown in Figure 17. To evaluate the time-domain performance of the vision sensor, displacements measured at point 9 and 16 by vision sensor are compared with those by LDSs. As shown in Figure 18, excellent agreements are observed between the two sensors, with the NRMS errors of 1.8% and 1.2%, respectively.

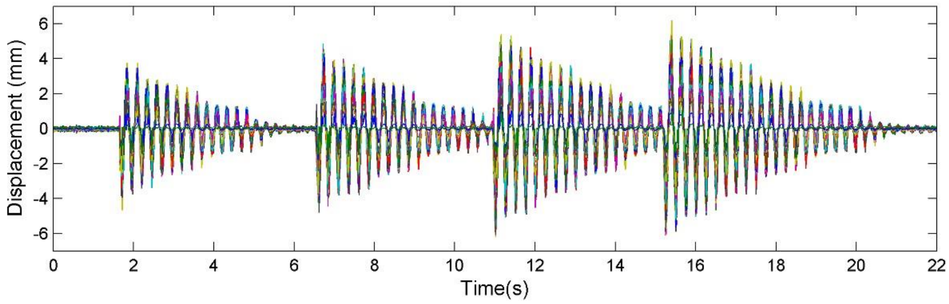


Figure 17. Displacement measurements at point 2 through 31 by vision sensor

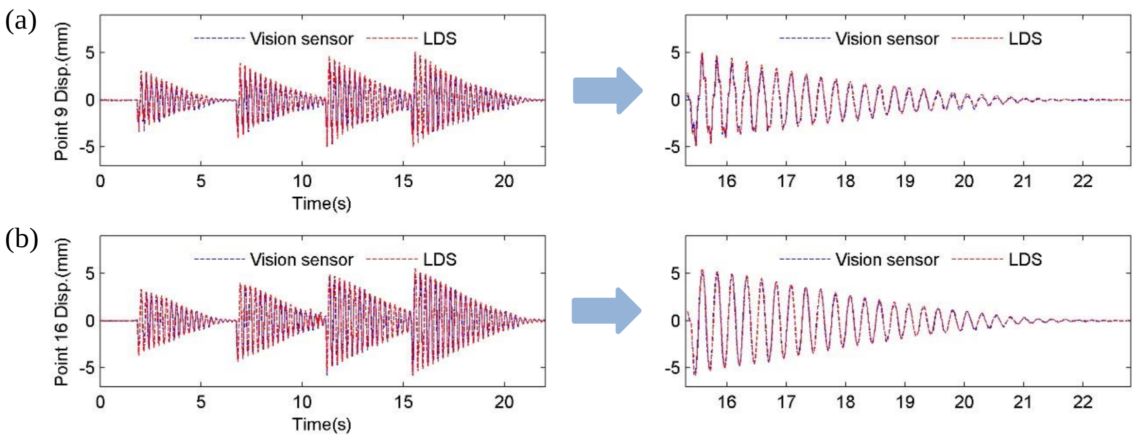


Figure18. Comparison of displacement measurement: (a) at point 9, and (b) at point 16

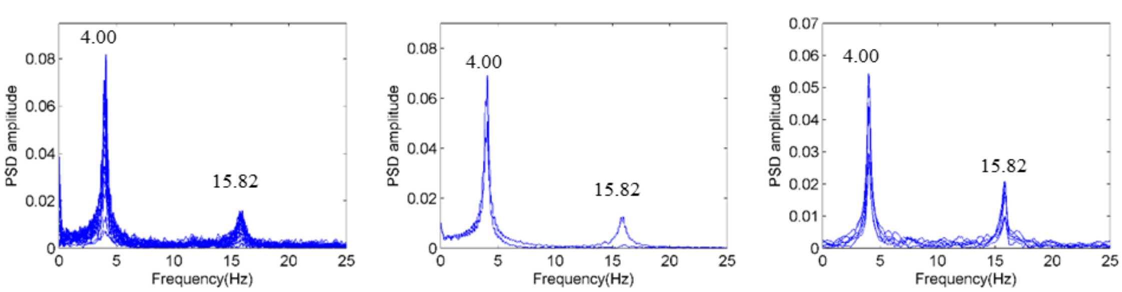


Figure19. Frequency results from: (a) displacement measurements at point 2 through 31 by vision sensor, (b) displacement measurements at points 9 and 16 by LDS, (c) accelerations at 6 points by

accelerometers.

Figure 19 compares the extracted natural frequencies by vision sensor, LDS and accelerometer, respectively, which shows a perfect match. Subsequently, the mode shapes of the simply supported beam model can be obtained from modal analysis. Figure 20 compares the first two mode shapes (scaled to 1) by accelerometers and vision sensor. It is indicated that the vision sensor can achieve smoother mode shapes while the resolution of mode shapes from accelerometers is limited by the sensor number. Moreover, compared with accelerometers, the vision sensor system is far more convenient and cost-effective for experimental modal analysis, since the displacements at multiple points can be simultaneously measured by a single camera placed at a convenient remote location, without the need for installing sensors on the structure.

(a)
(b)

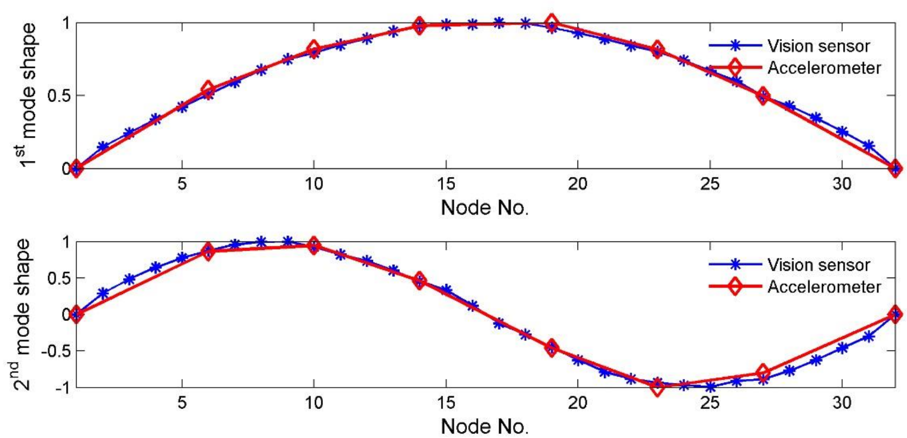


Figure 20. Comparison of mode shapes between vision sensor and accelerometer: (a) 1st mode shape, (b) 2nd mode shape

3.3.6 Damage Detection via MSC-Based Index

The use of the multipoint dynamic displacements simultaneously measured by a single camera for damage detection is demonstrated. Two identical simply support beam models, as shown in Figure 21(a), are tested. One beam has no damage while the other has two saw-cut dents to simulate damage with a 20% stiffness reduction, The conventional Mode shape curvature (MSC)-based damage detection index proposed by Pandey et al. (1991) is adopted. The MSC-based method is based on the premise that a reduction of stiffness associated with damage will cause an increase in the mode shape curvature, which can be calculated by the second-order central difference method. Originally, one of the major bottlenecks limiting the application of this technique is the resolution of identified mode shapes, which is associated with the number of sensors used. The dense full-field displacement measurement by the vision sensor makes it possible to construct much smoother mode shapes, and thus it is expected that the damage information can be revealed by the changes of the MSC or a normalized MSC (NMSC) between the damaged and undamaged structure. Based on the identified first two mode shapes from the vision sensor measurements, the MSC or NMSC damage index is calculated and plotted as shown in Fig.22. The beam damage can be obviously localized from the peak in the MSC or MMSC index using one set of vision sensor system.

This test demonstrates the significant potential of the vision-based displacement sensor system for a low-cost high-performance vibration-based structural health monitoring applications, such as experimental modal analysis and structural damage detection, compared with the conventional accelerometers.

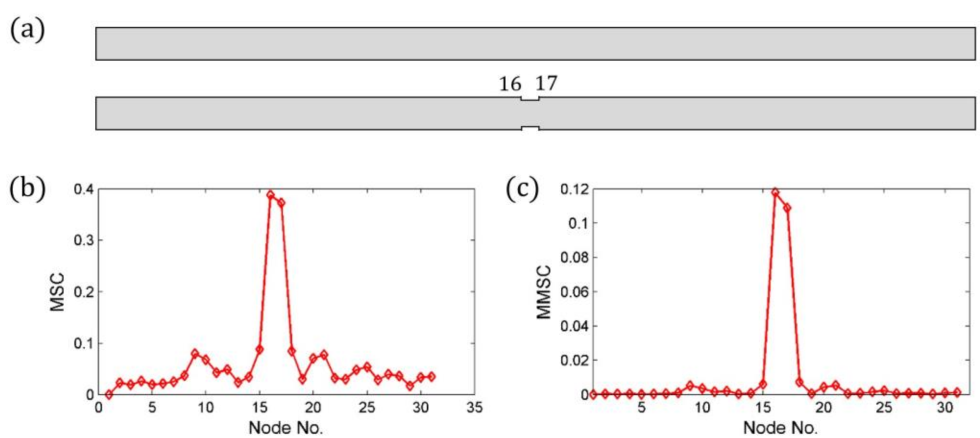


Figure 22. Damage detection results: (a) intact and damaged beams, (b) damage detection results by the MSC index, and (c) damage detection results by the MMSC index

3.4 Field Evaluation Tests at the Manhattan Bridge

It is highly difficult and expensive to use conventional sensors to measure the displacement responses of large-span bridges. Field tests are carried out at the Manhattan Bridge to study the performance of the vision sensor system for remote real-time multipoint displacement measurements. Manhattan Bridge, opened to traffic in 1909, is a suspension bridge spanning the East River in New York City and connecting Manhattan and Brooklyn. The main span is 448 meters long. The bridge is 36.5 meters wide, including 7 lanes in total and 4 subway lanes as shown in Figure 23.

The vision sensor system is placed on stable stone steps with negligible ground motions, around 300 meters away from the bridge mid-span position. For this long-period long-span bridge, a frame rate of 10 fps is used. The known dimension of the vertical truss members (7.2m, as shown in Figure 24 (a)) is used to estimate the scaling factors from SF_1 in Eq.(1). Note that the maximum frame rate of most low-cost video cameras is usually limited to 30~60 fps at the best resolutions. Such frame rates would be sufficient for measurement of most bridges that have natural frequencies of interest below 15 Hz. High-speed camera has limitations in practical applications because the higher the frame rate, the more difficult to achieve real-time processing of multipoint measurments.

(a) (b)

Figure 23. Manhattan Bridge: (a) test setup, and (b) cross-section

(a) (b)

Figure 24. Tracking target on the bridge: (a) one target, and (b) simultaneous measurements of 3 targets

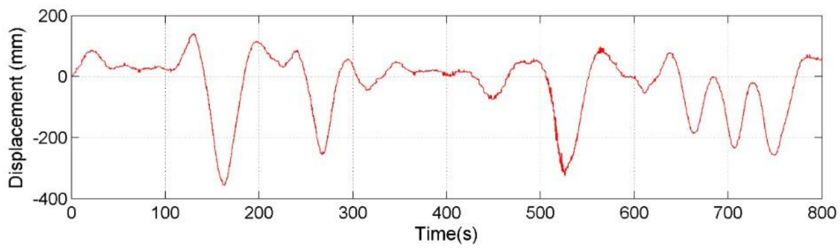


Figure 25. Displacement measurements of one target

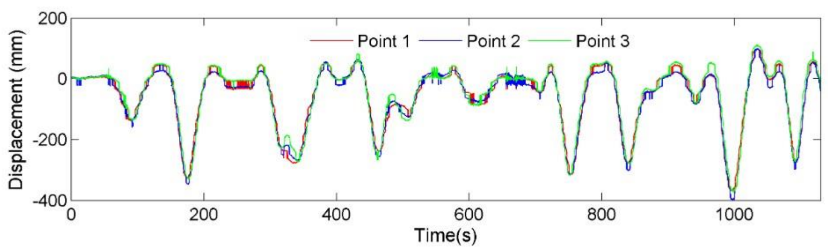


Figure 26. Simultaneous displacement measurements of 3 targets

Firstly, displacement responses at one single point at the mid-span region shown in 24 (a) are measured during the passage of subway trains. The scaling factor is estimated as 20.5mm/pixel. The dynamic displacement response recorded for a total of 800 seconds is plotted in Figure 25, the peak displacements (on the order of 300) of which are similar to those measured by both the GPS and interferometric radar systems in literature (Mayer et al, 2010). The response to the subway train can be clearly distinguished, with a maximum amplitude of 356mm. Then, by zooming out the lens to obtain a large field of view (FOV), i.e., area that is visible in the image, three points at the mid-span region are selected, as shown in 24 (b), and corresponding measurements are plotted in Fig. 26. Compared with Fig. 25, Figure 26 displays more fluctuations, especially when displacement amplitudes are small. This is because the scaling factor is estimated as about

36mm/pixel for this larger FOV case, meaning a decreased measurement resolution compared to the previous single point case. In practical applications, appropriate camera lens should be chosen so that its FOV is suitable for the specific tests.

Although the sensor accuracy is significantly improved through sub-pixel registration techniques, it is still limited by the camera hardware parameters, e.g., camera sensor resolution and the FOV, etc., especially for long-distance measurement of large-scale structures. In general, two measures can be taken to further improve the measurement accuracy: (1) increasing the resolution of the camera sensor (CCD or CMOS), which is a direct method to improve the accuracy when the measured area is fixed; and (2) utilizing an image magnification element to zoom in, as a result the measured area would accordingly decrease. Thus, tradeoffs between the measurement resolution and the FOV are necessary. To simultaneously measure full-field displacements along a large-scale structure such as a long-span bridge, multiple synchronized cameras can be used by targeting different sections of the structure.

4. HEAT HAZE FILTERING TECHNIQUE

The accuracy of remote displacement measurements in the field by a vision sensor depends on the image quality. Heat haze, also known as optical turbulence, is a complicated optical phenomenon which naturally occurs when viewing objects through a layer of heated air. The density variations in the air due to heat cause changes in the optical refraction index, resulting in distortions in the video images and measurement errors in the displacement obtained from the distorted images.

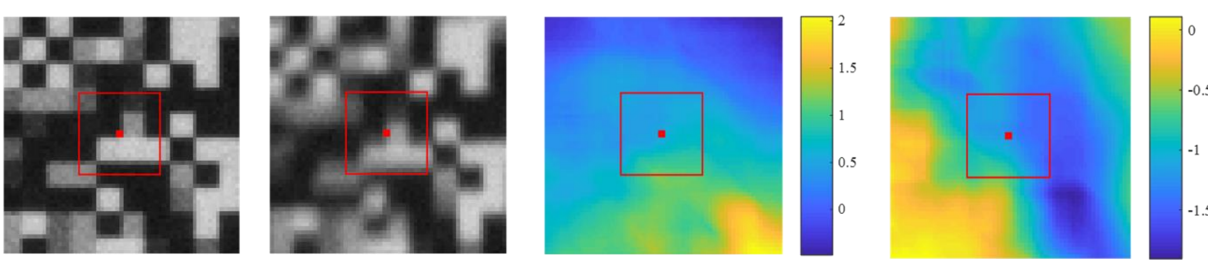
The heat haze problem has not been widely studied and current research mainly focuses on the techniques related to heat haze distortions in photographs involving stationary objects. These studies can be classified into three categories. The first category of researches provides techniques for recovering one good frame from aligning multiple affected frames through lucky imaging, image fusion, optical flow based super resolution, and space invariant deconvolution. These techniques, however, are designed for recovering stationary images, thus cannot distinguish structural displacements from the fluctuations caused by the haze. Moreover, aligning a series of distorted images is very time consuming, which limits its application in real-time monitoring. The second category of research work estimates distortion induced by heat haze by comparing the distorted/undistorted frames through optical flow, nearest neighbors, and statistical sampling. These techniques cannot differentiate the heat haze induced image distortions from the structural vibration. The third category investigates and quantifies the statistics of the heat haze distortion, but do not apply the statistics of heat haze distortion to remove the effect of heat haze in displacement measurement. Overall, the techniques developed in previous literatures are for stationary objects, and have not been applied to displacement measurement of moving objects.

This represents the first study of the heat haze problem in computer vision based displacement measurement. This study first investigates the effect of heat haze on the displacement measurement accuracy through laboratory and field experiments, and then analyzes the statistical pattern of the heat haze induced image distortions and measurement errors, based on which heat haze detection and filtering techniques are developed to minimize heat haze induced errors in displacement measurement. Finally, the heat haze detection and filtering techniques are validated in the laboratory experiments, and further demonstrated in the field tests conducted on the Williamsburg Bridge.

4.1 Measurement of Heat Haze inducted Image Distortions

The OCM template matching algorithm is further developed for 2D measurement of image distortions caused by heat haze. the surface displacement of each pixel is estimated usually within a selected region of interest (ROI). A sliding window with 21×21 pixels is place around a pixel and is tracked using the OCM template matching technique. The measured displacement is assigned to the center pixel of the sliding window as its image distortion. Similarly, the image distortions of all pixels within the ROI are measured by moving the sliding window pixel by pixel. The 2D measurement technique is applied to measure the heat haze induced image distortion.

An example is presented in Fig. 27 to illustrate the technique for image distortion measurement. Only the images and results within the ROI are shown. At first, in the reference frame, a sliding window (red frame) with 21×21 pixels is selected around a chosen pixel (red dot). Then the sliding window is tracked in the distorted frame, and its 2D location change is recorded as the image distortion of the center pixel in both axes. The image distortions of the other pixels in the frame are estimated similarly by moving the sliding window.



(a) (b) (c) (d)

Figure 27. Image distortion Measurement: (a) Reference frame, (b) Distorted frame, (c) Estimated vertical distortion, (d) Estimated horizontal distortion (unit: pixel).

4.2 Field Measurement of Heat Haze Induced Errors

Field tests are conducted to demonstrate the effect of heat haze on the measurement error of the vision sensor in outdoor natural environment. The field testing results are also used to study the statistical characteristics of the measurement errors. in field measurement.

Two field tests are conducted at the same location but in different months: first test on May 19th while the second on August 31st, 2017, both in the early afternoon. The weather condition is sunny with little wind. The camera is fixed on a solid ground 2,300 $meters$ away from the measurement target - a rigid building, as shown in Fig. 28. The long measurement distance results in thick layer of air, making the heat haze problem more prominent. The building is considered not vibrating in the vertical direction and the camera is isolated from ground ambient vibration and wind. The length of each test is approximately 30 minutes. Sampling rate of the tests was 60 frames per second (60 Hz). Since the building and camera are both considered stationary, any displacements obtained by tracking a target on the building can be considered as measurement errors due to heat haze.

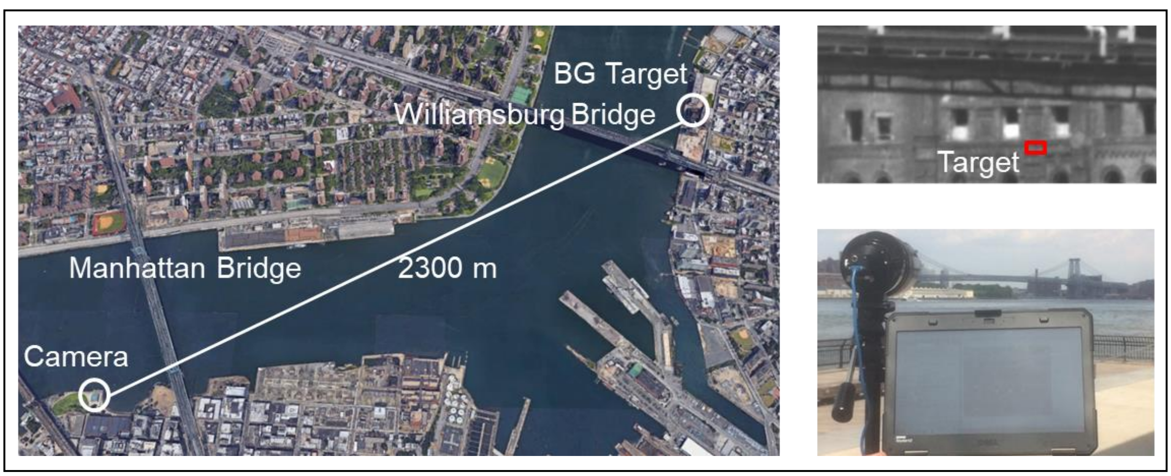
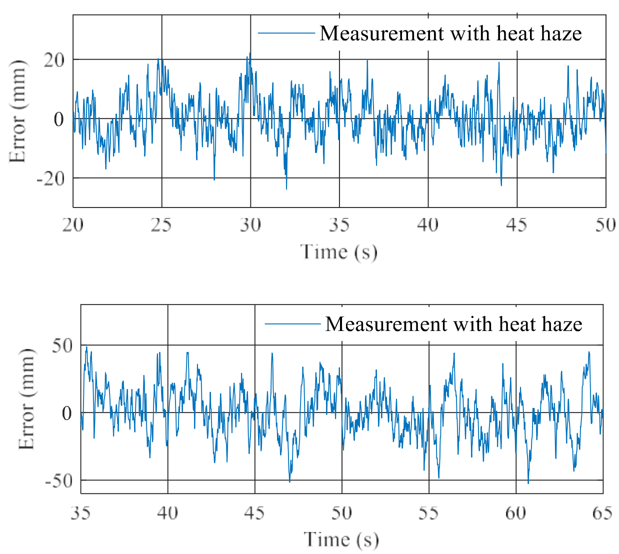


Figure 28. Setup of the field test for heat haze study.

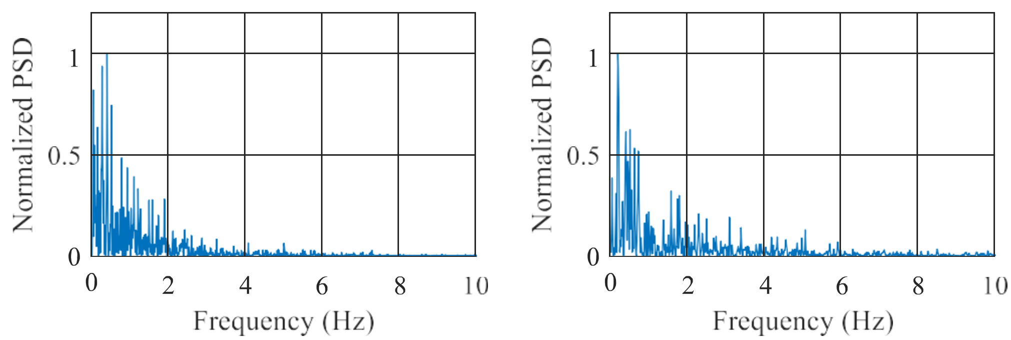
The measured video images show visible fluctuations caused by heat haze. The measured building vertical displacement in the two field tests are plotted in Fig. 29. RMSE of the measurements from two field tests is 16.44 mm and 7.24 mm. The difference in the RMSE of two field tests is probably due to different temperature and humidity conditions at the measurement time. The maximum magnitude of the measured vertical displacement reaches as high as 50 mm. Again, neither the building nor the camera is moving in the vertical direction and the measured displacements are considered as the measurement errors caused by the heat haze.

Frequency domain analysis is further performed on the field measurement results, and the power spectral density (PSD) functions of the heat haze induced measurement errors are plotted in Fig. 30. The dominant frequency is mostly in the low frequency range, overlapping with dominant frequencies of most large structures. Therefore, low pass or high pass filtering techniques cannot eliminate the heat haze induced measurement errors in the vision based displacement sensor. It is necessary to develop a effective filtering technique to mitigate the heat haze induced measurement errors in hot weather.



(a) First field test (b) Second field test

Figure 29. Measurement error due to heat haze in the field tests



(a) First field test (b) Second field test

Figure 30. PSD functions of the heat haze induced measurement error in the field tests.

4.3 Statistical Characteristics of the Heat Haze Distortions

In order to develop an effective filtering technique, this project further conducts laboratory and field experiments to study statistical characteristics of the image distortions and displacement measurement errors caused by heat haze. The analyses results inform the design of a heat haze detection and filtering system.

4.3.1 Laboratory Tests of Heat Haze

To further study the heat haze induced distortions, heat haze experiments are conducted in the laboratory to supplement the field measurement results. The experiments also provide additional datasets for validating the heat haze detection and filtering techniques.

Figure 31 shows the setup of the experiment. Dark heathers are used to produce heat haze. The camera is placed close to 12 $feet$ away from the target panel attached to a four-story frame model.

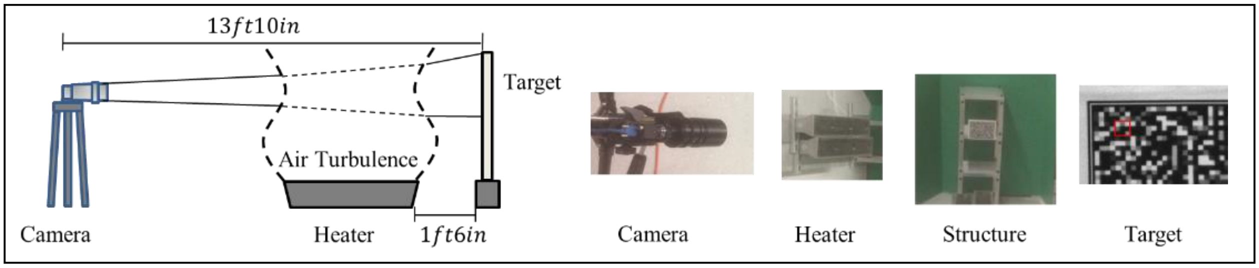


Figure 31. Setup of the laboratory experiment for heat haze.

Distortions in video images are results of spatial variations in the refraction indexes caused by non-uniformity of the densities in the heated air. Figure 32 shows an example of a frame of distorted images including its magnitude and phase, from which it is observed that the heat haze induced distortion is randomly distributed.

The distortion magnitudes of consecutive frames, two of which are presented in Fig. 33, show that the distortion distribution constantly shifts. Therefore, if the structural displacement is measured by tracking the target, the measurement can easily be affected by the heat haze induced image distortion. It was also found from Figs. 32 and 33 that some areas in a given distorted frame are less distorted than others, thus the displacement obtained from the least distorted targets would contain least measurement errors.

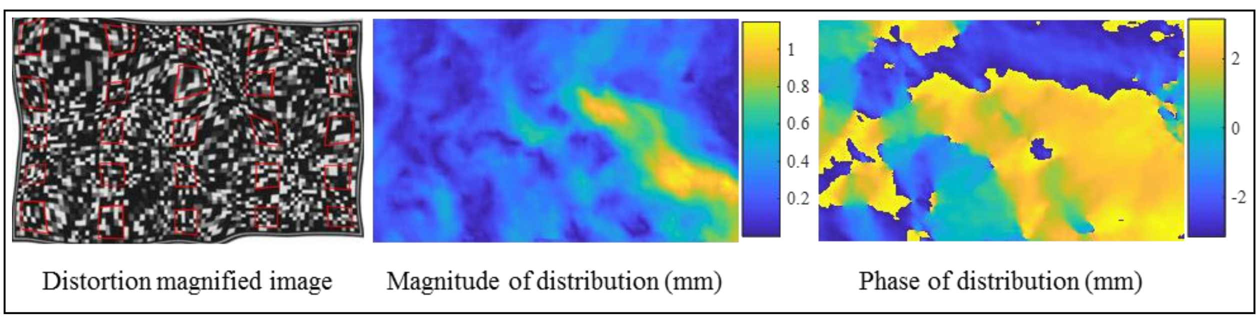


Figure 32. Estimation of heat haze distortion.

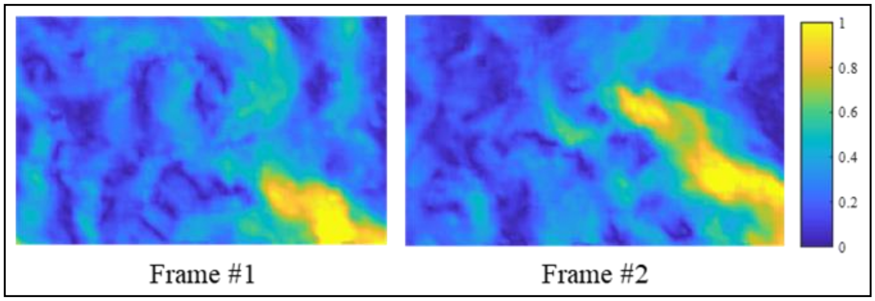


Figure 33. Change of the magnitude of heat haze distortion.

4.3.2 Statistical Characteristics of the Heat Haze Distortions and Field Validation

In order to detect and filter the heat haze induced image distortions, statistical characteristics of the measurement errors obtained from the distorted video images are analyzed. The heat haze induced displacement measurement errors can be assumed to follow the normal distribution, and the variance of the distribution is a function of the measurement distance and the diameter of the camera aperture:

$$\sigma_v^2 = 2.914 D^{-1/3} \int_0^L C_n^2 \left(\frac{x}{L}\right)^{5/3} dx = \frac{3}{8} \cdot 2.914 D^{-1/3} C_n^2 L \quad (3)$$

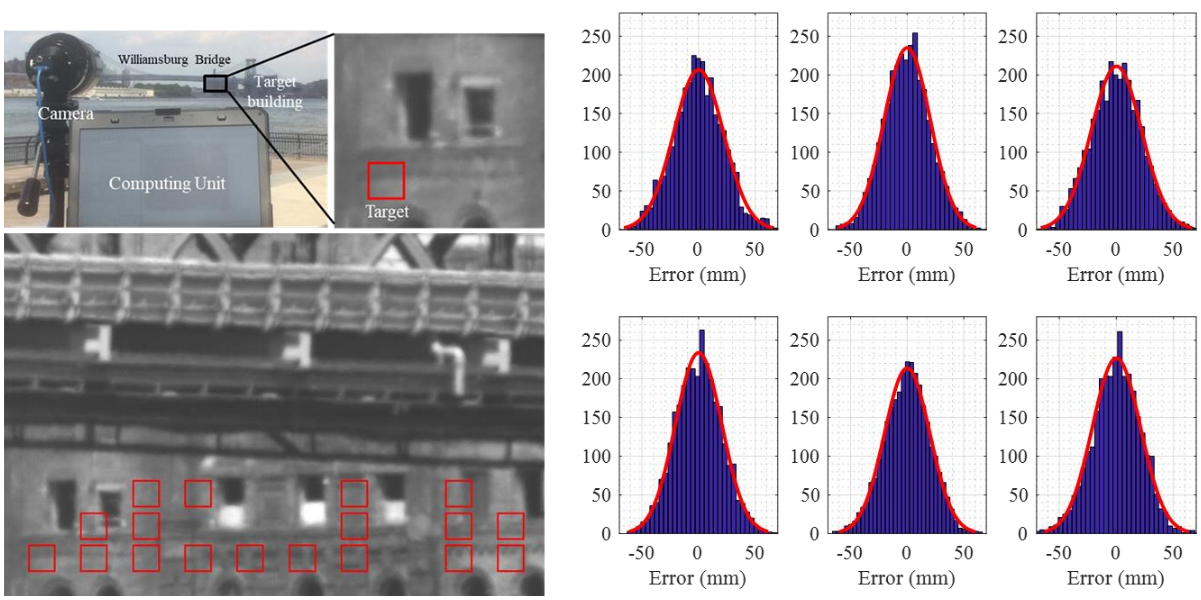
where σ_v^2 is the variance of the normal distribution that fits the heat haze error; D is the diameter of the aperture of the camera; L is the distance between the vision sensor and the object structure; C_n^2 is a variable that depends on the environmental and physical factors, such as temperature, pressure, humidity, etc.

The values of D and L are usually constant and the value of C_n^2 do not change suddenly in assumption, and thus the variance of the normal distribution is constant during a short period of time. Assume the errors are independent and identically distributed, the variance of the normal distribution that makes the observed errors the most probable given the distribution is the unbiased estimator of the variance of the observed errors.

$$\sigma_v^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})^2; \ where\ \bar{\epsilon} = \frac{1}{n} \sum_{i=1}^{n} \epsilon_i \quad (4)$$

More field tests, similar to the two tests presented in Section 4.2, are conducted to validate the statistical characteristics and investigate the effect of other factors such as target location and size. In the field tests, the targets are selected on a stationary structure so the results were not affected by structural vibration, and the camera is fixed on a solid foundation isolated from ambient and wind vibration. Therefore, the measured displacement by the vision sensor can be considered mainly due to heat haze induced image distortions.

First, the effect of the target location on the statistical characteristics of the heat haze errors is studied. Multiple targets are selected on the reference frame and their displacements were extracted from the distorted video images. The histograms of the heat haze errors of multiple targets are displayed in Figure , where normal distributions are fitted onto the histograms. As shown in the plots, the variances of the normal distributions fitting the errors of different targets have similar values. The results validate that the measurement errors can be fitted by the normal distribution and the target location does not have noticeable effect on the variance of the distribution.

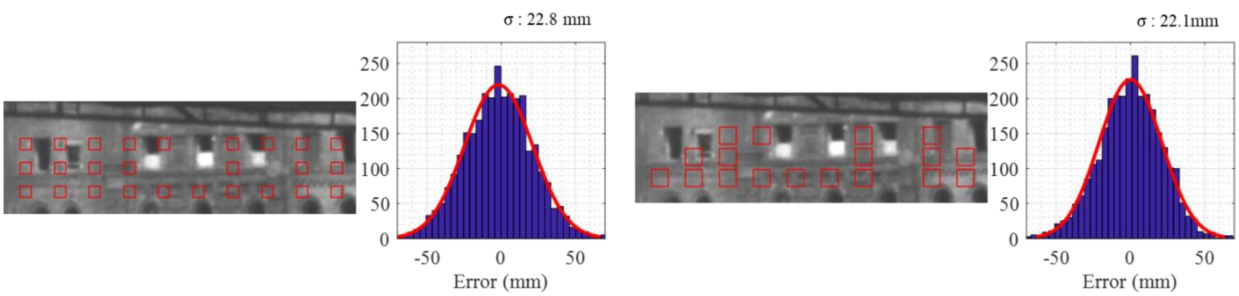


(a) Field test setup (b) Histogram of heat haze error and normal distribution fitting

Figure 34. Heat haze error obtained from different target locations in the field.

Second, the effect of the target size on the variance of the normal distribution is investigated. Targets with two different sizes (40×40 and 60×60) at the same location are tracked. The histograms of the heat haze errors obtained from tracking these targets are plotted in Fig. 35. Normal distributions are fitted onto the histograms of the measurement errors. It is observed that the variance of the normal distribution fitting the heat haze errors is not affected by the target size.

The results in the field tests validate that the variance of the normal distribution fitting the heat haze errors is not affected by the target location or target size. Since the physical factors that determine the value of C_T^2 are also assumed constant during a test, it is safe to conclude from the results in field tests that the heat haze induced errors can be fitted by the normal distribution and the variance of the distribution does not change during a test.



(a) Target of 40×40 pixels (b) Target of 60×60 pixels

Figure 1. Heat haze error obtained from different target sizes in the field.

4.4 Statistical Heat Haze Detection and Filtering System

Based on the statistical characteristics of the heat haze induced image distortions, a statistical heat haze detection and filtering system is developed in this project for the computer vision based displacement sensor. Figure 26 illustrates the overall system configuration, which involves two techniques. The first is to detect image distortions caused by heat haze. The triggering mechanism is enabled by a heat haze detection technique where a classifier is trained based on the features extracted from the distortion distributions of labeled images, some of which are affected by heat haze. When the classifier detects the effect of heat haze, the statistical heat haze filtering technique is constructed and applied.

The statistical heat haze filtering technique is composed of two steps and is constructed based on simultaneous multi-target displacement measurement and the statistical patterns of the heat haze induced distortions and measurement errors. In the first step, multiple targets are tracked, and the targets with the least distortion are identified and tracked to obtain the primary displacement. In the second step, a filter constructed based on the statistical characteristics of the measurement errors is applied to the primary displacement to further alleviate the heat haze induced measurement errors.

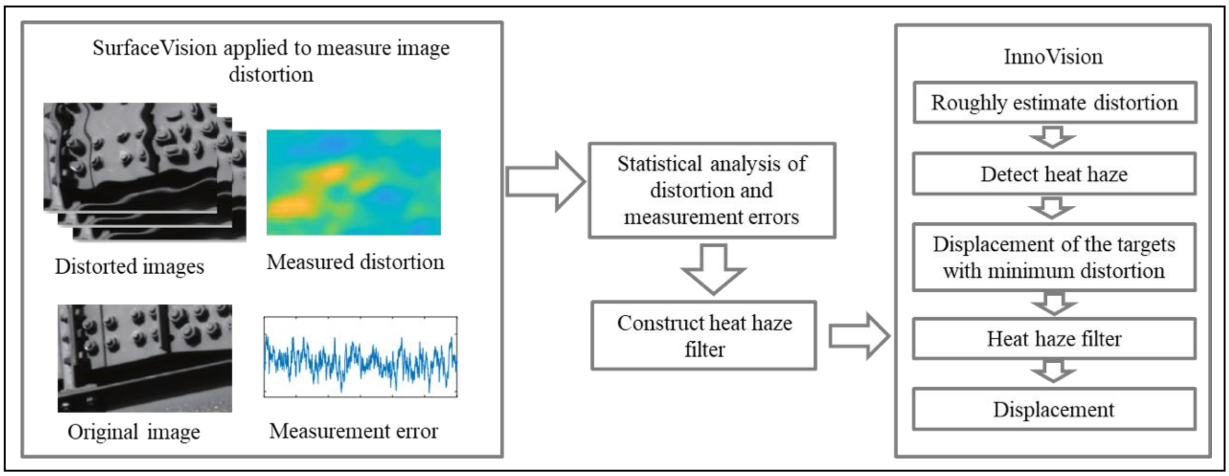


Figure 36. Overview of the heat haze detection and filtering system.

4.4.1 Heat Haze Detection Technique

The heat haze detection technique is proposed to detect heat haze induced image distortion in order to trigger the heat haze filter. The image distortion distribution on the video images is first estimated. Analyzing the distortion measurement, it is found that the heat haze significantly increases the spatial variance of the distortion distribution of a given frame. Therefore, a classifier can be trained to classify whether a new frame is distorted by heat haze, by applying a machine learning classification technique and the spatial features of the training samples. Frames unaffected and affected by heat haze are labeled as negative and positive training samples respectively, and the features used for training the classifier are the spatial variances of the distortion distributions in both axes of the training samples. The Naïve Bayes Classification algorithm is adopted in this study, and other advanced methods can also be used such as deep learning algorithms.

As an example, 150 positive samples and 150 negative samples are collected from the laboratory experiment, and as small percentage of the training samples can be randomly selected as hold out samples for testing the classifier. The classifier is trained using the Naïve Bayes classification algorithm and is represented in Figure 37 where the color defines the boundary between the normal frames not affected by heat haze and those affected by heat haze. The features $(f1,\ f2)$ represent the spatial features of the distortion distribution of a sample in both axes. The logarithm values of the features are used for classification. The obtained classifier can be applied to an unlabeled new frame to determine whether it is affected by the heat haze induced distortion.

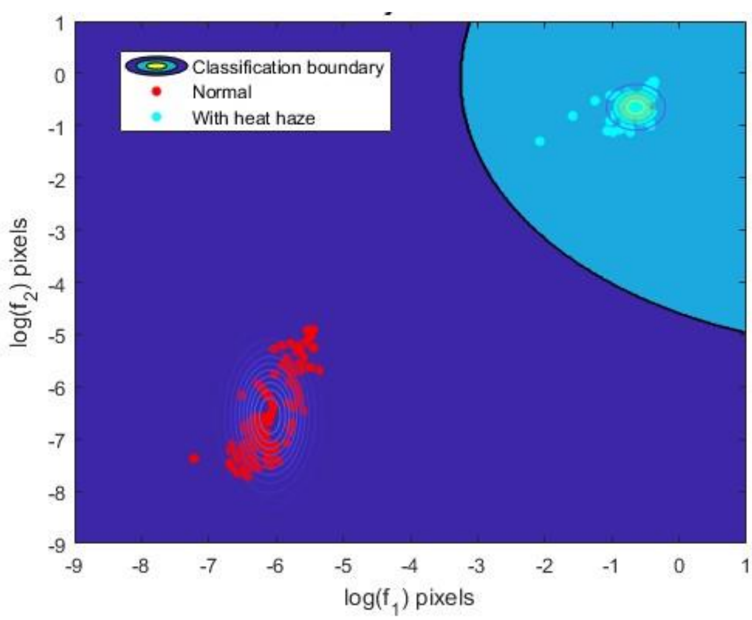


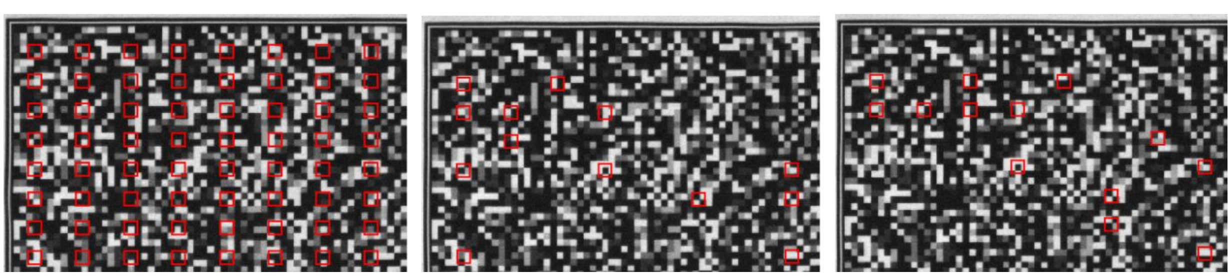
Figure 37. Classifier trained for heat haze detection.

4.4.2 Heat Haze Filtering Technique

Once heat haze is detected, the statistical heat haze filter is triggered. The filtering technique is composed of two steps. The first step is based on simultaneous multi-target tracking. Multiple targets are automatically selected on a frame and their displacements are tracked. By identifying and tracking the targets with the least image distortions, the primary displacement measurement is obtained. In the second step, a filter, with parameters determined based on the analysis of the statistical characteristics of the heat haze induced measurement errors, is applied to the primary displacement to further alleviate the errors.

As discussed earlier, the distortion distribution of a heat haze affected frame is random and constantly shifting. It is also found that some parts in the distorted frame are less distorted than the others, hence the displacements obtained from the least distorted targets have less measurement errors than that obtained from the other targets or one fixed target. Therefore, in the first step of the filtering technique, multiple targets (m targets) are selected within a region of interest (ROI) on the reference frame, and the distortions of these targets are estimated on the distorted frames. The distortion is estimated by calculating the root mean squared difference between the target in the current frame and that in the reference frame. Then, n targets ($n < m$) with the least distortion on each frame are identified and their displacements are tracked and averaged to obtain the primary displacements.

An example to illustrate the first step of the heat haze filtering technique is presented in Fig. 38 where only the images within the ROI are shown. Within the ROI of the reference frame, 8 by 8 targets were selected automatically. Then in any of the distorted frames, thirteen targets with the least distortion were identified and tracked for obtaining the primary displacement for the corresponding frame.



(a) Reference frame (b) Distorted frame #1 (c) Distorted frame #2

Figure 38. Demonstration of the first step of the filtering technique.

The second step of the statistical heat haze filtering technique is designed based on the statistical characteristics of the heat haze errors. As presented in previously, heat haze induced errors can be fitted by normal distributions and the variances of the distributions do not change. The primary displacement obtained from the first step of the filtering technique still contain heat haze errors in normal distribution errors, and the second step of the heat haze filtering technique is designed to further alleviate the measurement errors in the primary displacement. The filter is revised from the adaptive Wiener filter, and built based on the variance of the heat haze errors v which can be estimated from a segment of the data with little or no structural vibration. The errors are then filtered using a sliding window moving point by point.

In the revised filter, the mean μ and the variance σ^2 are calculated for the data within the sliding window $(i - M: i + M)$ that is applied from the beginning of the primary displacement:

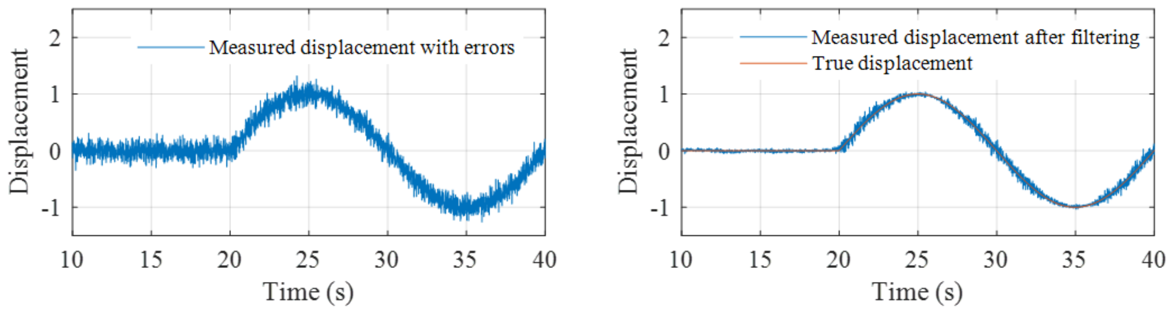
$$\mu = \frac{1}{2M+1} \sum_{n=i-M}^{i+M} x(n) \tag{4}$$

$$\sigma^2 = \frac{1}{2M+1} \sum_{n=i-M}^{i+M} (x(n) - \mu)^2 \tag{5}$$

Then the data in the center of the sliding window $x(i)$ is updated to $x^v(i)$ by deducting the normal distribution error from the primary displacement based on the variance of the errors . The measurement errors for all the data in primary displacement can be alleviated by moving the sliding window point by point.

$$x^v(i) = \mu + (1 - f(\sigma^2, v^2))(x(i) - \mu) \tag{6}$$

A simulated example is presented in Fig. 39 to show the effectiveness of the abovementioned filter. The measured displacements are mixed with heat haze induced measurement errors created from normal distribution. The variance of the measurement errors v is estimated from the first 20 seconds of the measured displacement in Fig. 39(a). Then the heat haze filter formulated based on the measured v is applied to the displacement, and the filtered displacement is compared with the true displacement in Fig. 39 (b). The heat haze errors subjected to the normal distribution is significantly reduced after applying the second step of the filtering technique.



(a) Measurement with errors (b) Measurement after filtering

Figure 39. Demonstration of the second step of the filtering technique.

4.5 Experimental Validation of the Heat Haze Detection and Filtering system

Laboratory experiment is conducted to validate the developed heat haze detection and filtering system using the three-story frame model and the conventional LVDT's as reference displacement sensors. The dynamic displacement measured by the vision sensor with and without the statistical filtering are compared with the reference LVDT data. A field test is further conducted on the Williamsburg Bridge to validate the efficacy of the heat haze detection and filtering techniques, by comparing the measured frequencies with analytical ones.

4.5.1 Laboratory Validation of the Heat Haze Detection and Filtering System

The setup of the laboratory experiment is presented in Fig. 40. Dark heaters generates heat haze. The three-story frame structural model is fixed on a shaking table and subjected to sinusoidal excitations.

Firstly, the heat haze detection technique was tested on two unlabeled frames where the first frame was obtained without heat haze and the other frame was obtained after the air was heated by the dark heaters. The features of the image frames (the spatial variances of the measured distortions) are extracted. Then the trained heat haze classifier is applied to determine whether the new frames are affected by heat haze. As shown in the classification results in Figure, the heat haze classifier successfully classified the first frame as a normal frame and the second frame as a heat haze affected frame. The classification results validated the efficacy of the heat haze detection technique.

Secondly, the statistical heat haze filtering technique was tested in the laboratory experiment. LVDT was used as the reference sensor since the laser displacement sensor relies on light reflection and might be vulnerable to air turbulence. The displacements and the measurement errors before/after applying the heat haze filtering technique are plotted in Figure and compared with the reference data. Before applying the heat haze filtering technique, the displacement measurement was erroneous due to the heat haze induced distortion. After the heat haze filtering technique was applied, the accuracy of the displacement measurement was increased and the filtered displacement matched well with the reference data. From analysis of RMSE of the displacement measurements, the measurement error was significantly reduced by about 67.5% from $0.0845\ mm$ to $0.0275\ mm$ after the statistical heat haze filtering technique was applied. The experiment results validated the efficiency of the statistical heat haze filtering technique for reducing heat haze errors in computer vision based displacement measurement.

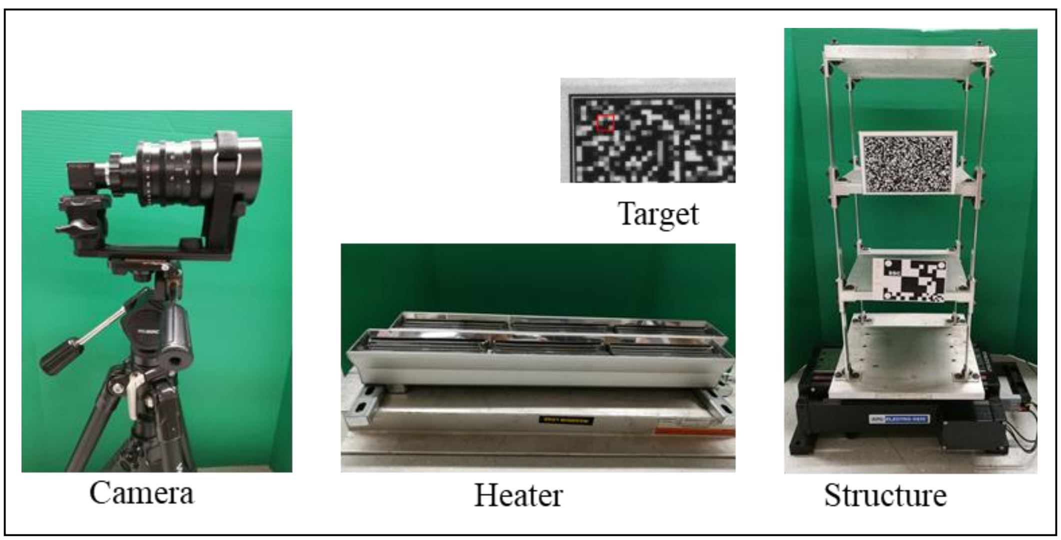


Figure 40. Experiment for validating the heat haze detection and filtering system.

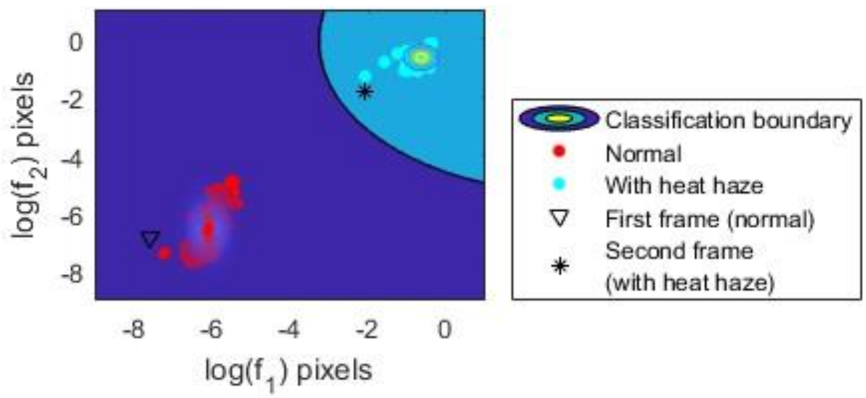
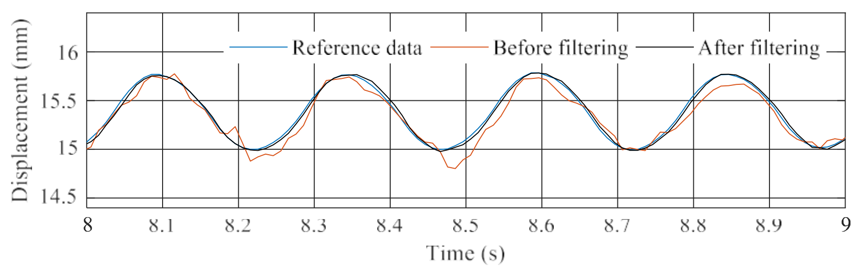
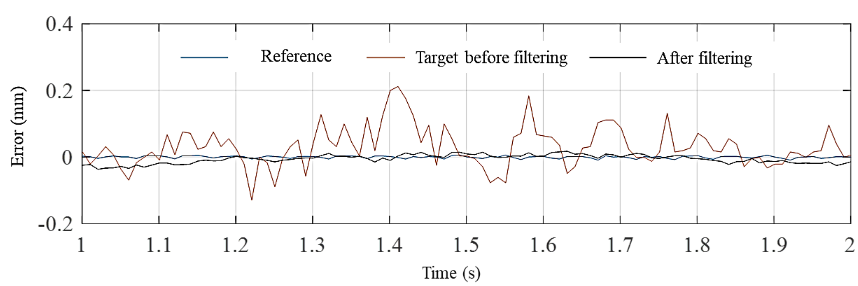


Figure 41. Classification result in the laboratory experiment



(a) Displacement measurement



(b) Measurement error

Figure 42. Measurement result from laboratory experiment before/after heat haze filtering

4.5.2 Filed Validation of the Heat Haze Detection and Filtering System

To further demonstrate the statistical heat haze filtering technique in the field environment in hot weather, a field test is conducted on the Williamsburg Bridge, a 2,227 $meter$ long subway train/highway bridge in New York City. A natural target is selected near the mid-span of the bridge for tracking the vertical dynamic response of the bridge. As shown in Fig. 43, the camera is located at about 2,080 $meter$ away from the target and a long distance telescopic lens was used.

The test is conducted on May 19th, 2017, when the mean temperature is 26 ℃ and the maximum temperature is 33 ℃ on the day of the test. The vision sensor images are visibly affected by the heat haze. While the camera is fixed on the ground, a practical camera vibration cancellation technique (to be presented in Chapter 5) is applied using a stationary background target (BG target) on the stationary building behind the bridge. The camera vibration is cancelled by tracking the BG target and the bridge target simultaneously and subtracting the BG displacement from the bridge target displacement.

The raw displacement time history and the results after camera vibration cancellation and heat haze filtering are plotted in Fig. 44. From the video footage, it is found that that a subway train passes through the bridge between 50 to 100 *sec*. But the passing train load is not correctly reflected in the raw displacement time history. After applying the vibration cancellation technique, a reasonable displacement response is revealed. Then the two-step heat haze filtering is applied. In the first step, thirty targets are tracked for each frame, which five targets with the least distortions are identified and their displacements are tracked and averaged to obtain the primary displacement. Subsequently, the heat haze filtering technique step two is applied to further alleviate the heat haze errors. After applying the heat haze filtering, the displacement time history becomes more smooth.

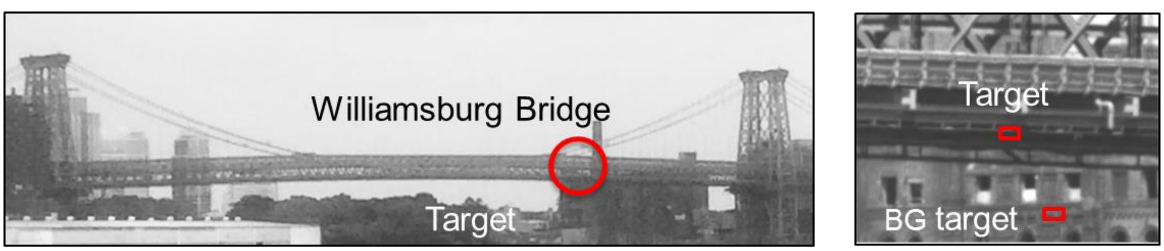


Figure 43. Field test for demonstrating the statistical heat haze filtering technique.

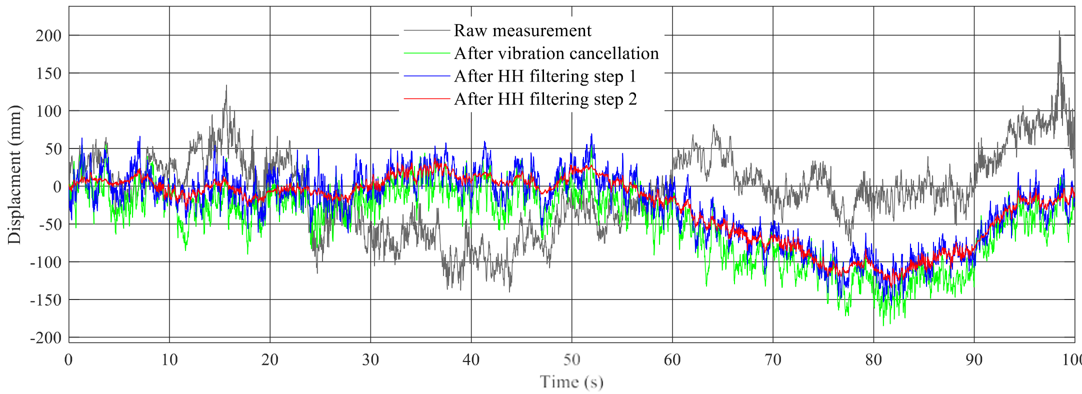
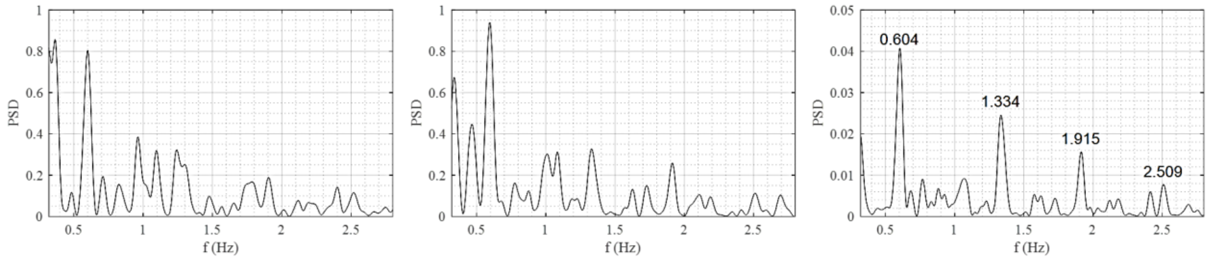


Figure 44. Displacement obtained from field test before/after heat haze filtering

Frequency domain analysis of the displacement measurement is conducted to validate the measurement results. The PSD of the raw displacement and the displacements after camera vibration cancellation and heat haze filtering are plotted in Fig. 45. After applying the techniques, the PSD becomes smoother and the resonant frequencies can be clearly identified. The first, second, third, and fourth resonant frequencies of the bridge are 0.604 Hz, 1.334 Hz, 1.915 Hz, and 2.509 Hz respectively. They agree well with the theoretical values computed as follows.

As stated in the literature (Feng et al. 2015), when a train with multiple carriages passes through, resonant frequencies of the bridge response under the train load can be obtained as: $f = nV/L;\ n = 1;\ 2;\ 3;\ \ldots$ where V is the speed of the train and L is the distance between two carriage centers. For the M-line or the Z-line subway train which pass through the Williamsburg Bridge, the R32 or the R42 car models are used. The distance between two carriage centers is approximately 18 meters and the total length of the train is approximately 160 meters. From the video footage, it takes 15 seconds to pass the bridge measurement point over the entire length of the train, thus the train speed is estimated as approximately 10.7 m/s. Therefore, the theoretical value of the first resonant frequency of the bridge response is #`.Ä Å/5 / #a Å = 0.594 Hz, which agrees very well with the first resonant frequency identified from the PSD plot. The rest of the theoretical resonant frequencies: 0.189 Hz, 1.783 Hz, 2.378 Hz also satisfactorily agree with the measured values. The results of the field test demonstrated the efficacy of the heat haze filtering technique in the field environment.



(a) From raw measurement (b) After vibration cancellation (c) After heat haze filtering

Figure 45. Power spectrum density of measurement results.

5. CAMERA VIBRATION CANCELLATION TECHNIQUE

Camera vibration caused the inevitable ground ambient vibration and wind is the another challenge to measurement accuracy when using vision sensors in the field. The effects of camera vibration become more significant when the target being monitored is located far away from the camera. A practical technique for camera vibration cancellation is proposed in this study by applying the subpixel enhanced multipoint measurement method discussed in Chapter 2.

In the proposed technique, the object targets, referred to as the targets, are selected on the monitored structure, together with the background target, referred to as the BG target, on the background object that is considered stationary. The displacements of the targets and the BG target are measured simultaneously. Because the BG target is considered stationary, any measured BG target displacement would be the measurement error due to camera vibration. Therefore, to eliminate such error, the BG target displacement should be subtracted from the background target. Laboratory and field experiments are carried out to validate this technique.

5.1 Laboratory Validation

The experimental setup is shown in Fig. 46. The camera vibration is simulated by placing the camera on a shaking table, which is excited by white noise signal. Two targets on the floors of a two-story frame and one BG target at an adjacent stationary structure are selected and their displacements are tracked simultaneously. Two laser displacement sensors (LDS) are installed on the stationary structure, as reference sensors. The video camera is located 4.2 meters away from the target structure and the SF of video images is $0.7724\ \mathrm{mm/pixel}$. The test is conducted in the low light condition.

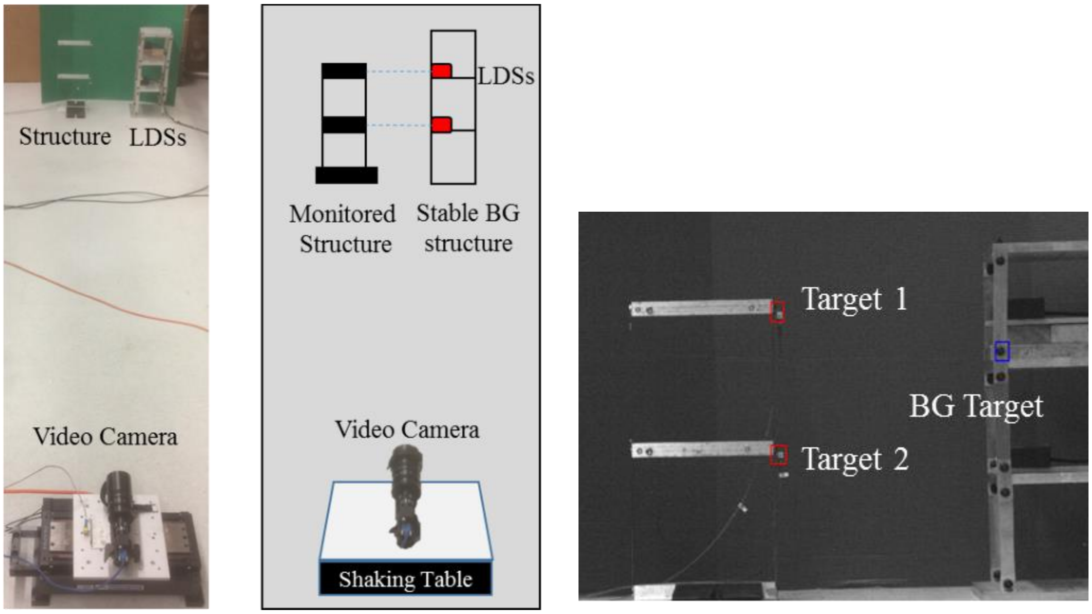
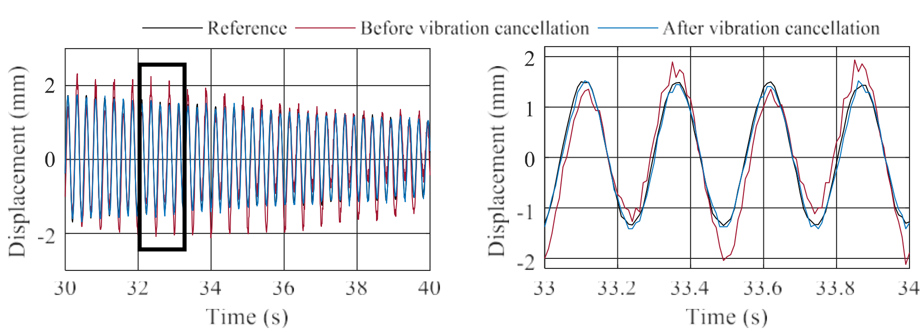
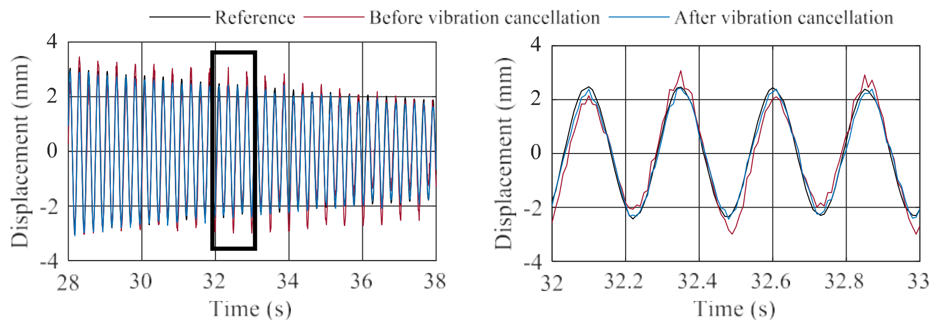


Figure 46. Experiment setup and target selection.

A hammer is used to induce impact force on the second floor of the testing frame. The floor displacements of the frame due to the impact force are recorded by the reference LDSs as well as the vision sensor. Figure 47 shows the displacement time histories measured by the vision sensor before and after camera vibration cancellation, in comparison with those measured by the reference LDSs. The displacements measured by the vision sensor after the camera vibration cancellation agree very well with those by the reference LDSs. Using the proposed camera vibration cancellation technique, the RMSE errors are reduced by 61%. (from 0.409 mm to 0.158 mm) for the second floor and by 53.81% (from 0.446 mm to 0.206 mm). The laboratory results validates the efficacy of the practical camera vibration cancellation technique based on the simultaneous multipoint displacement measurement.



(a) Displacement of floor 1



(b) Displacement of floor 2

Figure 47. Displacement measurement before/after camera vibration cancellation

5.2 Field Validation at the Manhattan Bridge

The camera vibration cancellation technique is further evaluated in the field, in which the vision sensor measures the mid-span vertical displacement of the Manhattan Bridge, a long-span suspension bridge as described earlier. As shown in Fig. 48, the camera is on purposely placed on the Brooklyn Bridge, located 447 *meters* away from the measurement point, in order to emphasize the camera vibration effects. The scaling factor for the bridge is estimated at 26.76 mm/pixel. The sampling rate is 60 Hz.

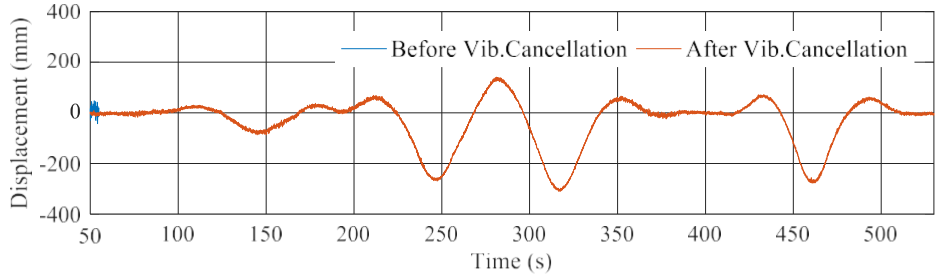
A natural target is selected at the mid-span of the Manhattan Bridge and a BG target on a background building. The building is considered stationary in the vertical direction. The displacements of the target and the BG target are simultaneously measured by the vision sensor. The camera vibration is canceled by subtracting the displacement of the BG target from the displacement of the target. The sampling rate of the displacement measurement was 60 Hz. Since there is not a stationary platform for installing high fidelity displacement sensor such as LVDT or laser displacement sensor, the measurement result was validated through dynamic analysis. Acceleration data taken at the mid-span of the bridge were used as reference data to compare with the measurement results in the frequency domain.

(a) Bird-view of testing site (b) Test setup

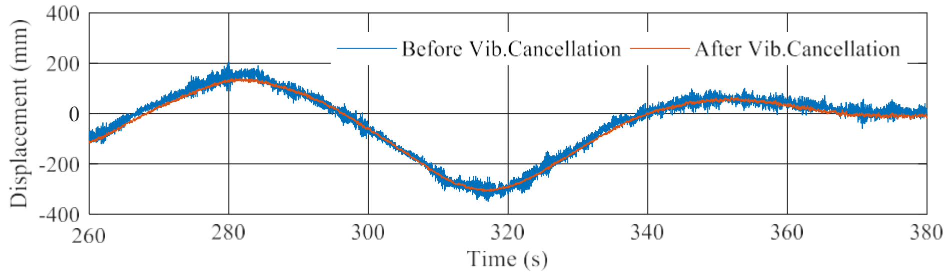
Figure 48. Field test bird-view and setup.

Figure 49 shows the subway train induced bridge displacement time histories measured with and without the camera vibration cancellation technique applied. Their PSDs are plotted in Fig. 50, showing the difference that the camera vibration cancellation is making.

The results are compared with those obtained from previous dynamic tests on the bridge using accelerometers and GPSs. Table 2 compares the resonant frequencies obtained from the previous tests with those from this tests, in which V1, V2 and V3 represent the first three vertical resonant frequencies and T1 represents the first torsional resonant frequency. After applying the vibration cancellation technique, the first four vertical mode frequencies are identified as 0.22 Hz, 0.30 Hz, 0.40 Hz, and 0.50 Hz, which agree very well with those obtained from the previous tests. Without cancelling the camera vibration, the resonant frequencies identified from the displacement measurement do not match the reference data. This validates the necessity and the effectiveness of the camera vibration cancellation technique developed in this study.

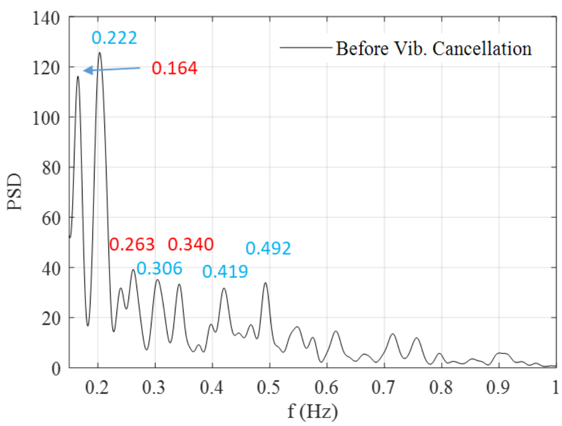


(a) Long displacement plot for 500 seconds

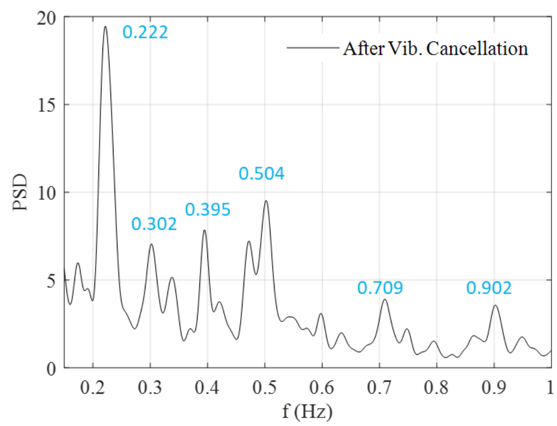


(b) Short displacement plot for 120 seconds

Figure 49. Displacement measured before/after camera vibration cancellation



(a) PSD before vibration cancellation



(b) PSD after vibration cancellation

Figure 50. Frequency plot before/after camera vibration cancellation

Table 2. Resonant frequencies before/after camera vibration cancellation

Data type	Data Year	1st resonant V1 (Hz)	2nd resonant V2 (Hz)	3rd resonant T1 (Hz)	4th resonant V3 (Hz)

Reference acceleration data [37]	2009	0.23	0.30	0.37	0.50
Displacement before vibration cancellation	2017	0.16	0.22	0.26	0.31
Displacement after vibration cancellation	2017	0.22	0.30	0.40	0.50
Recent acceleration data	2016-2017	0.23	-	0.40	0.51
GPS data [38]	Before 2009	0.23	-	0.30	0.49

5.3 Field Validation at the Jamboree Bridge

The vision sensor system is further evaluated on a short-span concrete bridge, the Jamboree Road Overcrossing, in Irvine, CA. The main span of this box-girder bridge is 60.1 *meters*. The camera is fixed on the ground near the freeway underneath the bridge, and the freeway traffic causes camera vibration. A target is selected on the bridge mid-span and a BG target on a building, as shown in Fig. 51. The vision sensor measures the vertical displacements of the target and the BG target and subtracts the latter from the former.

Figure 51. Setup for the second field test

Figure 52 shows an example of the measured target displacement of the bridge (displacement before vibration cancellation), the BG target displacement (camera vibration), and the bridge displacement after applying the camera vibration cancellation technique. The raw vertical displacement of the bridge (without camera vibration cancellation) shows a similar drifting trend to the BG target displacement, indicating that most of the target displacement is caused by the camera vibration. Without cancelling the camera vibration, the vision sensor would give misleading measurement results. This field test again demonstrates the importance of considering the effect of the camera vibration on the measurement when using the vision sensor in the field. Applying the vibration cancellation technique, the bridge displacement time history becomes much more reasonable, demonstrating the effectiveness of the proposed practical camera vibration cancellation technique.

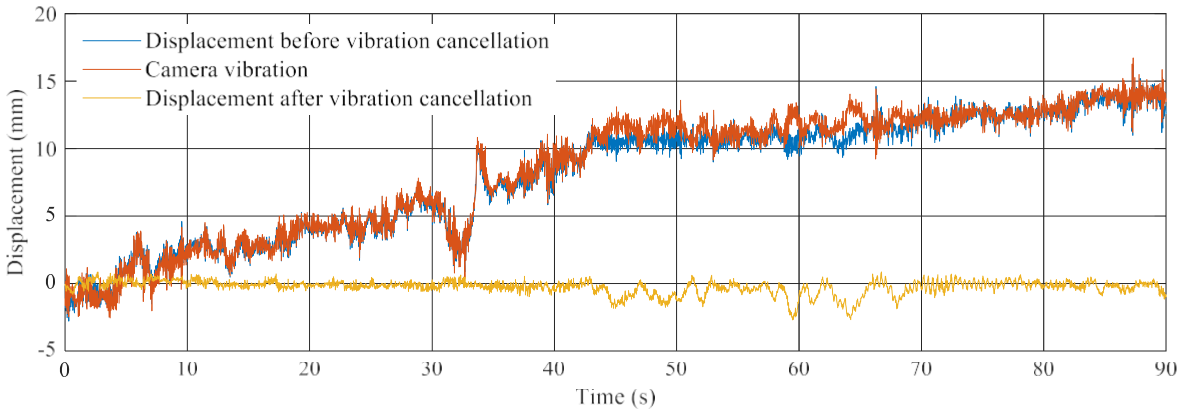


Figure 52. Measurement results in field tests of Jamboree Bridge

6. CONCLUSIONS

Rapid advances in computer vision promise low-cost, easy-to-operate sensors for convenient measurement of bridge dynamic responses at a remote location without interrupting the bridge operation. This project addresses challenging issues associated with field application of the computer vision sensor technology including robust tracking of multiple natural targets in ill environments and effects of heat haze and camera vibration. A number of vision sensing techniques are developed including (1) a gradient-based OCM template matching algorithm incorporated with a subpixel technique to increase spatial resolution, enabling multipoint measurement, (2) a statistical heat haze detection and filtering system, and (3) a multipoint measurement-based technique for camera vibration cancellation. These algorithms are integrated into a software package and extensively evaluated through laboratory and field tests. Major findings and conclusions are listed as follows.

1. The OCM template matching algorithm enhanced by the subpixel technique enables accurate measurement of structural displacements at multiple points using a single camera. From the shaking table tests of a frame and a beam structural models, excellent agreements are observed between the displacement time histories measured at multiple points by the vision sensor via tracking low-contrast natural targets on the structures and those by high-fidelity reference laser displacement sensors.

2. Investigations of the robustness of the vision sensor indicate that, in case of challenging conditions such as illumination fluctuation, partial template occlusion, or background disturbance which are frequently encountered in field tests, the gradient-based OCM template matching algorithm is more reliable than the conventional intensity-based UCC algorithm, since the employed orientation codes are inherently invariant to variations in the image intensity.

3. This study recommends to estimate the scaling/calibration factor based on the known physical dimension on the object surface and the corresponding image dimension in pixels. For the cases of non-perpendicular lens optical axis, scaling factors in the horizontal and vertical

directions should be estimated respectively. By this method, satisfactory measurement accuracy can be guaranteed for small camera tilt angles.

4. The remote, real-time and multipoint measurement capabilities of the vision sensor system developed in this study are further validated through field tests on the Manhattan Bridge during train passing, with the vision sensor system placed more than 300 meters away from the bridge. However, due to the limits in the camera resolution and the subpixel technique, tradeoffs between the measurement resolution and measurement points or field of vision is necessary. To measure displacements over the entire deck of a long-span bridge, several synchronized cameras can be employed.

5. This study represents the first effort in studying the heat haze effects on displacement measurement errors. Using the multipoint 2D template matching techniques developed in this study, numerous laboratory and field experiments are conducted to characterize the patterns of heat haze-induced image distortions, which is found random and constant shifting. Subsequently, heat haze detection and filtering techniques are developed based on the statistical characteristics. The effectiveness of these techniques are demonstrated through extensive laboratory and field tests involving long-distance measurements in hot weather conditions.

6. A practical solution is developed to cancel the camera vibration by simultaneously tracking the displacements of both the monitored structure and a stationary reference point, then subtracting the displacement of the reference point from the structural displacements. The effectiveness of this solution is validated through laboratory tests involving effects of camera view angles as well as the field tests on the long-span Manhattan Bridge and the short-span Jamboree Bridge.

7. This project further demonstrates the unique advantages of the vision sensor for low-cost structural health monitoring and damage detection applications. Damage of a beam structural model, together with its location, is successfully detected, thanks to the simultaneous measurement of displacements of 30 points, which enables construction of a high-resolution mode shape curvature index and accurate damage localization. This would be costly and

difficult to achieve by conventional accelerometers, because 30 of them would need to be installed on the structure.

PUBLICATIONS

This project has produced the following published journal and conference papers:

Feng, D., & Feng, M. Q. (2017). Experimental validation of cost-effective vision-based structural health monitoring. *Mechanical Systems and Signal Processing*, 88, 199-211.

Feng, M. Q., Fukuda, Y., Feng, D., & Mizuta, M. (2015). Nontarget vision sensor for remote measurement of bridge dynamic response. *Journal of Bridge Engineering*, 20(12), 04015023.

Luo, L., & Feng, M. Q. (2017). *Vision based displacement sensor with heat haze filtering capability*. *Proc. of the International Workshop on Structural Health Monitoring,* Stanford University, CA.

Luo, L., & Feng, M. Q. (2018). Heat Haze Detection and Filtering for Computer Vision Based Displacement Measurement. *Mechanical Systems and Signal Processing*, Manuscript submitted for publication.

Luo, L., Feng, M. Q., & Wu, Z. Y. (2018). Robust vision sensor for multi-point displacement monitoring of bridges in the field. *Engineering Structures*, 163, 255-266. doi:https://doi.org/10.1016/j.engstruct.2018.02.014

ACKNOWLEDGEMENT

This project is supported by National This project is supported by the Cooperative Highway Research Program (NCHRP) Highway program under the contract Grant No. 20-30/IDEA 189. We would like to express our sincere gratitude to Dr. Inam Jawed, Program Manager, for his dedicated guidance and support. We are also grateful to the Advisors of this project, Ben Foster and Dale Peabody of Maine DOT, for their invaluable advice and discussions throughout the entire project. The project has received in-kind support from the New York

City DOT and the California Department of Transportation for the bridge field evaluation tests.

REFERENCES

Baqersad J, Poozesh P, Niezrecki C, Avitabile P. Photogrammetry and optical methods in structural dynamics - A review. Mechanical Systems and Signal Processing. 2016.

Busca G, Cigada A, Mazzoleni P, Zappa E. Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System. Experimental Mechanics. 2014;54:255-71.

Caetano E, Silva S, Bateira J. A vision system for vibration monitoring of civil engineering structures. Experimental Techniques. 2011;35:74-82.

Dworakowski Z, Kohut P, Gallina A, Holak K, Uhl T. Vision-based algorithms for damage detection and localization in structural health monitoring. Structural Control and Health Monitoring. 2016:35-50.

Feng, D.M. (2016) Advanced Vision-Based Displacement Sensors for Structural Health Monitoring, PhD Dissertation, Columbia University

Feng, D.M., Feng MQ, Ozer E, Fukuda Y. A vision-based sensor for noncontact structural displacement measurement. Sensors 2015;15(7):16557-16575.

Feng, D.M., Feng MQ. Vision-based multipoint displacement measurement for structural health monitoring. Structural Control and Health Monitoring 2016; 23(5):876-890.

Feng, D.M., Scarangello T, Feng MQ, Ye Q. Cable tension force estimate using novel noncontact vision-based sensor. Measurement: Journal of the International Measurement Confederation 2017;99:44-52.

Feng, D.M., Feng MQ. Identification of structural stiffness and excitation forces in time domain using noncontact vision-based displacement measurement. Journal of Sound and Vibration 2017;406:15-28.

Feng, D.M., Feng MQ. Experimental validation of cost-effective vision-based structural health monitoring. Mechanical Systems and Signal Processing 2017;88:199-211.

Feng, D.M., Feng MQ. Computer vision for SHM of civil infrastructure: From dynamic response measurement to damage detection – A review. Engineering Structures 2018;156:105-117.

Feng MQ*, Fukuda Y, Feng, D.M. Mizuta M. Nontarget vision sensor for remote measurement of bridge dynamic response. Journal of Bridge Engineering 2015;20(12).

Ho H-N, Lee J-H, Park Y-S, Lee J-J. A Synchronized Multipoint Vision-Based System for Displacement Measurement of Civil Infrastructures. The Scientific World Journal. 2012;2012:9.

Kim S-W, Jeon B-G, Kim N-S, Park J-C. Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge. Structural Health Monitoring. 2013;12:440-56.

Kohut P, Holak K, Uhl T, Ortyl Ł, Owerko T, Kuras P et al. Monitoring of a civil structure's state based on noncontact measurements. Structural Health Monitoring. 2013.

Lee J-H, Ho H-N, Shinozuka M, Lee J-J. An advanced vision-based system for real-time displacement measurement of high-rise buildings. Smart Materials and Structures. 2012;21:125019.

Lee JJ, Shinozuka M. A vision-based system for remote sensing of bridge displacement. NDT & E International. 2006;39:425-31.

Mayer L, Yanev B, Olson LD, Smyth A. Monitoring of manhattan bridge for vertical and torsional performance with GPS and interferometric radar systems. Transportation Research Board 89th Annual Metting. 2010.

Min JH, Gelo NJ, Jo H. Non-contact and Real-time Dynamic Displacement Monitoring using Smartphone Technologies. Journal of Life Cycle Reliability and Safety Engineering. 2015;4:40-51.

Olaszek P. Investigation of the dynamic characteristic of bridge structures using a computer vision method. Measurement. 1999;25:227-36.

Ozer E*, Feng, D.M., Feng MQ. Hybrid motion sensing and experimental modal analysis using collocated smartphone camera and accelerometers. Measurement Science and Technology 2017;28(10).

Pandey AK, Biswas M, Samman MM. Damage detection from changes in curvature mode shapes. Journal of Sound and Vibration. 1991;145:321-32.

Ribeiro D, Calçada R, Ferreira J, Martins T. Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. Engineering Structures. 2014;75:164-80.

Santos CA, Costa CO, Batista JP. Calibration methodology of a vision system for measuring the displacements of long-deck suspension bridges. Structural Control and Health Monitoring. 2012;19:385-404.

Song Y-Z, Bowen CR, Kim AH, Nassehi A, Padget J, Gathercole N. Virtual visual sensors and their application in structural health monitoring. Structural Health Monitoring. 2014;13:251-64.

Stephen GA, Brownjohn JMW, Taylor CA. Measurements of static and dynamic displacement from visual monitoring of the Humber Bridge. Engineering Structures. 1993;15:197-208.

Wahbeh AM, John PC, Sami FM. A vision-based approach for the direct measurement of displacements in vibrating systems. Smart Materials and Structures. 2003;12:785.

Wu L-J, Casciati F, Casciati S. Dynamic testing of a laboratory model via vision-based sensing. Engineering Structures. 2014;60:113-25.