Vibration Sensing of a Bridge Model Using a Multithread Active Vision System

Tadayoshi Aoyama, *Member, IEEE*, Liang Li, Mingjun Jiang, Kazuki Inoue, Takeshi Takaki, *Member, IEEE*, Idaku Ishii, *Member, IEEE*, Hua Yang, *Member, IEEE*, Chikako Umemoto, Hiroshi Matsuda, Makoto Chikaraishi, and Akimasa Fujiwara

***Abstract*—This study proposes a concept for multithread active vision sensing that can measure dynamically changing displacement and vibration at multiple points on civil engineering structures. In multithread active vision sensing, a high-speed camera can function virtually as multiple tracking cameras by accelerating its measurement, computation, and actuation with ultrafast viewpoint switching at millisecond level. This enables simultaneous measurement of small vibrations distributed across a wide range, which cannot be observed using a single camera system, because its pixel accuracy is generally incompatible with its wide angle of view. We developed a galvanomirror-based high-speed multithread active vision system that can switch 500 different views in a second; it functioned as 15 virtual cameras each operating at 33.3 fps to observe multiple scenes in completely different views. The experimental results for a 4-m-long truss-structure bridge model to which 15 markers were attached show that a single active vision system can observe the deformation of the bridge structure and estimate modal parameters, such as resonant frequencies and mode shapes, at a frequency on the order of dozens of hertz.**

***Index Terms*—Active sensing, high-speed vision, multithread viewpoint control, structure health monitoring, vibration analysis.**

Manuscript received February 9, 2017; revised June 21, 2017; accepted September 24, 2017. Date of publication October 19, 2017; date of current version February 14, 2018. Recommended by Technical Editor Y. Shen. This work was supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (KAKENHI) under Grant 16H02348. *(Corresponding author: Tadayoshi Aoyama.)*

T. Aoyama is with the Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Nagoya 464-8603, Japan (e-mail: tadayoshi.aoyama@mae.nagoya-u.ac.jp).

L. Li is with the Process Control System Software Design Department, Hitachi High-Technologies Corporation, Tokyo 105-8717, Japan (e-mail: liang.li.fg@hitachi-hightech.com).

M. Jiang, K. Inoue, T. Takaki, and I. Ishii are with the Department of System Cybernetics, Hiroshima University, Higashi-Hiroshima 739-8527, Japan (e-mail: akitoshi@hiroshima-u.ac.jp; k-inoue@robotics.hiroshima-u.ac.jp; takaki@robotics.hiroshima-u.ac.jp; iishii@robotics.hiroshima-u.ac.jp).

H. Yang is with the State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430073, China (e-mail: huayang@hust.edu.cn).

C. Umemoto is with Keisoku Research Consultant Company, Hiroshima 732-0029, Japan (e-mail: chikakoh@krcnet.co.jp).

H. Matsuda is with the Department of Structural Engineering, Nagasaki University, Nagasaki 852-8521, Japan (e-mail: matsuda@nagasakiu.ac.jp)

M. Chikaraishi and A. Fujiwara are with the Graduate School for International Development and Cooperation, Hiroshima University, Higashi-Hiroshima 739-8529, Japan (e-mail: chikaraishim@hiroshima-u.ac.jp; afujiw@hiroshima-u.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMECH.2017.2764504

I. Introduction

Vibration monitoring of civil engineering structures, including bridges and buildings, is an important technology for structural health monitoring to inspect the degrees of structural damage in their dynamic properties. Most vibration analyses to measure the dynamic responses of civil engineering structures have been conducted using contact-type sensors such as accelerometers or strain gauges [1]–[4], or noncontact-type optical sensors such as laser vibrometers or radar interferometry systems [5]–[8]. When installing contact-type sensors, the sensor must be attached on a reference point on the structure; however, there are many cases in which it is difficult or impossible to access the point on the structure on which it is to be installed. Owing to the acute directivity of laser beams, laser vibrometers can remotely measure vibration displacement with high accuracy. However, most optical sensors are designed for small displacement measurements at a single point by using high-intensity laser beams for long-distance measurement, which may endanger human eyes.

For noncontact dynamic displacement and vibration measurement featuring low cost and flexibility in its application to structures, many vision-based studies using commercial standard digital cameras operating at dozens of frames per second have been reported [9], [10] such as vibration and force measurement for cables [11], [12], digital-image-correlation-based displacement measurement for railway bridges [13], cantilever-beam measurement using subpixel Hough transform [14], rotational angle measurement for large civil structures [15], and dynamic displacement measurement of large structures with a robust object search algorithm [16]. Audio-frequency-level structural vibration analyses using high-frame-rate (HFR) videos have also been reported in [17]–[21], and recently several studies have used real-time HFR vision systems operating at hundreds or thousands of frames per second for simultaneous structural vibration analyses [22], [23].

Despite the fact that several digital-image-correlation methods have dealt with subpixel displacement measurement, the accuracies of vision-based approaches are still limited to the order of the pixel pitch of an image sensor. It is especially difficult for a single camera to simultaneously observe small vibration displacements at many points that are distributed across a wide area of a large-scale structure, because the total number of pixels of the image sensor and the optical resolution of the camera lens are insufficient to observe both the global shape of the large-scale structure and the local small vibrations in a single image. A multicamera network using multiple active pan-tilt-zoom (PTZ) cameras, which has been studied and developed actively in the field of security-based surveillance [24], [25], is one of the methods to observe real-world information in a wide area. This method requires as many active PTZ cameras as the number of measurement points to simultaneously observe small displacements at many points on a large-scale structure. However, as the number of PTZ cameras increases in a large-scale multicamera network that consists of dozens or hundreds of cameras, many issues, such as the physical size of multiple PTZ cameras and the complexities in camera control software, become more serious problems. If a single active camera system could simultaneously measure small displacements at many points distributed on a large-scale structure instead of multicamera-network-based surveillance, the installation and management cost of the vision-based measurement system could be remarkably reduced in vibration analysis for large-scale structures.

Therefore, in this study, we propose a concept of multithread active vision in which a single high-speed tracking camera can function virtually as multiple pan-tilt tracking cameras by accelerating its measurement, computation, and actuation with ultrafast viewpoint switching. We developed a galvanomirror-based high-speed multithread active vision system that can switch and process 500 different views in a second. These views cannot be simultaneously observed using a single camera. Several structural vibration analyses have been conducted for a truss-bridge model in which the multithread active vision functioned as 15 virtual cameras operating at 33.3 fps. The remainder of this paper is organized as follows. Section II introduces the concept of multithread active vision for large-scale-structure monitoring. Section III gives an outline of the galvanomirror-based multithread active vision system and its specification. Section IV shows its effectiveness for structural vibration analysis by presenting vibration measurement results for a 4-m-long truss bridge model, including estimation of modal parameters, such as resonant frequencies and mode shapes at dozens of hertz. Section V discusses the concerns that arise about the application of the proposed method to the measurement of structures that are dozens of meters long. In Section VI, we give the conclusion and discuss future work.

II. Multithread Active Vision Sensing

As mentioned in the introduction section, it is difficult to apply a multicamera network using multiple active PTZ cameras to vibration sensing for large-scale structures because of the physical size of multiple PTZ cameras or complexities in the control software. In addition, high-magnification lenses are required to obtain high-resolution images when we observe large-scale structures; thus, it is very difficult to set multiple positions and to calibrate multiple points by using different multiple cameras. Thus, a single camera system for vibration sensing is required for large-scale-structure analysis; however, it is difficult to realize simultaneous vibration observation at many points on a large-scale structure using a single camera system because its resolution is insufficient. If we observe 10- or 100-m-scale structures in a megapixel image, 1 pixel corresponds to 10 or 100 mm; it is difficult to measure small displacements of the order of 1 mm. In these circumstances, 10 000 $\times$ 10 000 pixel or 100 000 $\times$ 100 000 pixel image sensors are required to achieve 1-mm measurement accuracy. Currently, there are no such high-resolution image sensors; additionally, the optical resolution of the camera lens is insufficient. Therefore, this paper proposes a multithread active vision system that can perform as a virtual multi-PT (pan and tilt) camera using a single camera system. The multithread active vision system is also highly competitive in price because high-magnification lenses are quite expensive. In general, the resonant frequency of large-scale structures is low, that is, a few hertz. In this paper, the multithread active vision is realized by 15 virtual cameras operating at 33.3 fps. The 33.3-fps camera suffices for observation of large-scale structures. The main difficulties in visual observation of large-scale structures are the setting problem for a multiple camera system and the low resolution of a single camera system. The multithread active vision system overcame both of these problems. In this study, we assumed that the multithread active vision is applied under the condition that the subject distance is 100 m in scale and a high-magnification lens is used. Thus, the depth of focus is sufficiently long to obtain clear images of multiple viewpoints.

Recently, high-speed vision systems operating at 1000 fps or more have been developed [26], [27]. Object tracking systems, such as optical flow [28] or face tracking [29], are accelerated using these high-speed vision systems; the effectiveness of the high-speed vision system has been confirmed in robotic control [30], [31], flying-object tracking [32], and micro-object tracking [33]. Although imaging and processing of these vision-based tracking systems is accelerated, the actuator speed is insufficient; therefore, the tracking systems are based on visual feedback, which requires a few frames for convergence. Thus, the independent control of pan and tilt angles at every frame is difficult and the tracking system cannot function as a virtual multi-PT camera system.

The following conditions are required so that a single active vision system has a potency equivalent to that of multiple pan-tilt cameras.

1) Acceleration of imaging and processing: When each virtual PT camera is shooting at dozens of frames per second, the frame capturing and processing rate of an actual single vision system must perform the imaging and processing for multiple images of the virtual PT cameras. A rate of several hundreds or thousands of frames per second is necessary in order to realize dozens of virtual PT cameras.
2) Acceleration of viewpoint control: High-speed and independent viewpoint control must ensure that a frame does not affect the next frame in order to control the viewpoint at every frame corresponding to the acceleration of imaging and processing. Thus, viewpoint control via a

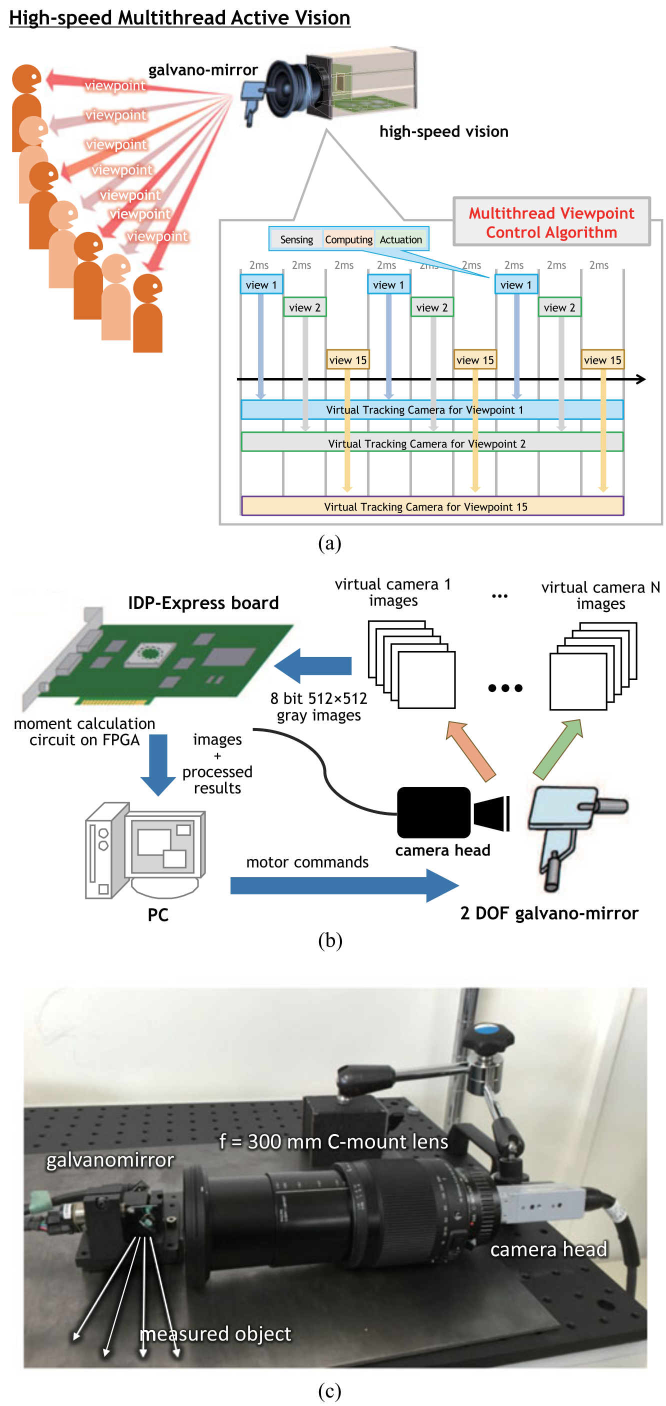


Fig. 1. Concept of the multithread active vision sensing. (a) Multithread viewpoint control. (b) Configuration of multithread active vision. (c) Overview of multithread active vision.

high-speed actuator that has a frequency characteristic of a few kHz is necessary.

Fig. 1(a) shows the concept of multithread viewpoint control based on meeting the two required conditions described above. The multithread viewpoint control is the concept that performs viewpoint control of virtual multi-PT cameras on a single active vision system by implementing time-division multithread processing of viewpoint control. Multithread processing is the concept that threads conducting tasks are simultaneously performed in a processor with time sharing. This idea of multithread viewpoint control is extended to PT camera control including imaging, processing, and viewpoint control. The multithread viewpoint control requires not only parallelization of processing but also minimization of the temporal granularity of the time-division thread processing time for imaging and viewpoint control. Thus, the imaging, processing, and viewpoint control, including the motion of the actuator, must be accelerated. In this paper, high-speed hardware system is considered in order to realize acceleration of the imaging, processing, and viewpoint control.

III. Galvanomirror-Based Multithread Active Vision System

We developed a high-speed multithread active vision system for real-time vibration analysis. In this paper, a high-speed vision system and a galvanomirror are used in order to accelerate image processing and viewpoint control based on the concept of the multithread active vision. The multithread active vision system consists of the high-speed vision platform IDP Express [34], a galvanomirror (6210H, Cambridge Technology), and a control personal computer (PC) (Windows 7 Enterprise 64-b OS, ASUS P7P55D-E board, Intel Core i5 760 2.8-GHz CPU, 4-GB DDR3 1333-Hz memory). A DA board (Interface, PEX-340416) to send control signals to the galvanomirror and an AD board (Interface, PEX-321216) to collect the pan and tilt angles of the galvanomirror are mounted on the control PC. Fig. 1(b) and (c) shows the system configuration and the system overview, respectively.

IDP Express consists of a compact camera head, the dimensions and weight of which are 23 mm $\times$ 23 mm $\times$ 77 mm and 145 g, respectively, when no lens is mounted, a dedicated FPGA board (IDP Express board), and a PC. The camera head can capture and transfer 8-b gray-level 512 $\times$ 512 images to the IDP Express board at 2000 fps. An $f = 300$ mm C mount lens (18–300 mm F3.5–6.3 DC MACRO OS HSM, Sigma) is attached to the camera head.

In order to accelerate image processing, the following calculation module is implemented in hardware to achieve high-speed calculation; it can output at a maximum rate of 2000 fps. In this study, we used simple image processing, that is, the binarization of the input image and calculation of the image centroid, described above, to obtain the position of the viewpoints in the measurement object. This simplification of the image processing allows its implementation in hardware for high-speed calculation and to realize robust calculation of the viewpoints' position.

A. Binarization of Input Image

The input image of the camera head captured at time t, $I(x, y, t)$, is converted to a binary image $B(x, y, t)$

$$B(x, y, t) = \begin{cases} 1, & (I(x, y, t) \leq \theta_B) \\ 0, & (\text{otherwise}) \end{cases} \tag{1}$$

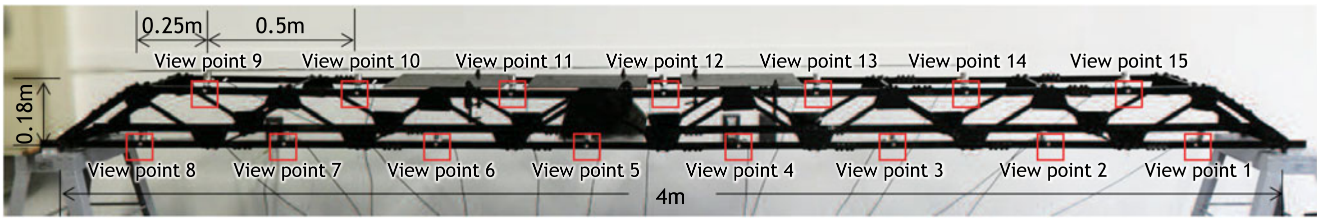


Fig. 2. Truss-structure bridge model.

where θ_B is the threshold.

B. Image Centroid Calculation

The zeroth-order and first-order moment features of $B(x,y,t)$ are computed as

$$M_0 = \sum_{x,y} B(x,y,t) \tag{2}$$

$$M_x = \sum_{x,y} xB(x,y,t),\ M_y = \sum_{x,y} yB(x,y,t). \tag{3}$$

The position of a tracked marker in the input image of the camera is calculated as the following image centroid:

$$(c_x, c_y) = \left(\frac{M_x}{M_0}, \frac{M_y}{M_0}\right). \tag{4}$$

The galvanomirror can control 2-DoF viewpoints based on pan and tilt angle mirrors. Both pan and tilt mirrors can control in the range of -20° to 20° by applying a voltage signal via the DA board mounted on the PC. The size of the pan and tilt mirrors are 17.5 mm^2 and 10.2 mm^2, respectively. We confirmed that the galvanomirror angle can be controlled within 1 ms if the mirror angle is in the range of 20°.

In this paper, we specified a 2-ms frame interval consisting of 1-ms viewpoint control and 1-ms exposure and processing time as the minimum thread unit time. Whether the number of measurement points increases or not, we consider using a minimum frame interval in order to realize as high a performance measurement as possible.

IV. Experiments

A. Bridge Model and Experimental Settings

The system performs one active vision as 15 virtual 33.3-fps cameras by using multithread viewpoints processing along an allocated visual line of 500 viewpoints/s for 15 virtual camera tasks as shown in Fig. 1(a).

The proposed multithread system is implemented based on time-sharing processing. After the position of the galvanomirror and the measurement object are fixed, the galvanomirror-based multithread active vision system does not need a special calibration or setup, because the position between the measurement object and the angles of the galvanomirror becomes a pair. In addition, in the preliminary experiment, we confirmed that 1 pixel of each viewpoint corresponds to 0.280–0.294 mm; the displacement difference of each viewpoint is small. Thus, we set the unit transformation from pixel to millimeter according to the average value of 15 viewpoints, that is, 1 pixel = 0.286 mm.

The experimental bridge model used in this research is a truss structure, as shown in Fig. 2. L-25 mm $\times$ 25 mm $\times$ 1.2 mm L-type aluminum members were used for up-and-down chord members and FB-15 mm $\times$ 2 mm flat aluminum members were used for bent and left-and-right chord members. The length, height, and depth of the bridge are 4, 0.18, and 0.3 m, respectively. The experiments were conducted at a condition in which 500 mm $\times$ 300 mm and 18.1-kg-weight steel plates were placed on the top of the bridge model in order to adjust the resonant frequency of the bridge model. The distance from the camera to the center of the bridge model was 5.5 m; 1 pixel corresponds to 0.286 mm, and the 512 $\times$ 512 image corresponds to an area of 146 mm $\times$ 146 mm. In total, 10 mm $\times$ 10 mm squared retroreflective markers were attached to the centers of 15 up-and-down chord members in the side of the bridge model. Each markers' position is derived based on the image centroid calculation described in Section III.

The vibration distribution of the 15 points on the bridge model were measured at six metal halide illumination conditions. Generally, acceleration sensors are used for health monitoring of civil structures and FFT analysis or modal analysis are conducted [37], [38]. In order to get comparative data, 15 piezoelectric pickups (PV-87, Lion) to measure vertical acceleration were attached to the same positions as the retroreflective markers; the vibration distribution was obtained using a vibration indicator unit (UV-15, Lion).

In this research, the viewpoint of the high-speed multithread active vision system was changed in the following order: the marker point $1 \rightarrow 2 \rightarrow \ldots \rightarrow 15 \rightarrow 1 \rightarrow \ldots$ in clockwise rotation. Viewpoints were changed every 2 ms; each marker was imaged every 30 ms. When the viewpoints moved along the trajectories $1 \rightarrow \ldots \rightarrow 8$ and $9 \rightarrow \ldots \rightarrow 15$, the pan angle was rotated by 4.64°. When the viewpoints moved from $8 \rightarrow 9$ and $15 \rightarrow 1$, the tilt angle was rotated by 1.66°. Fig. 3(a) and (b) shows the pan and tilt angle trajectories of the high-speed multithread active vision system. Fig. 3(c) and (d), respectively, shows the captured images and binarized images every 2 ms when the bridge model was static and number stickers were attached next to the markers on the bridge model to obtain the number of viewpoints. From Fig. 3, we can confirm that viewpoint movement finishes within 1 ms: the pan/tilt angles were standing for more than 1 ms in the frame interval of 2 ms and the viewpoint did not move in the exposure time of 1 ms. Also, it is confirmed that the pan/tilt angles were controlled to monitor the positions of the 15 markers attached to the bridge

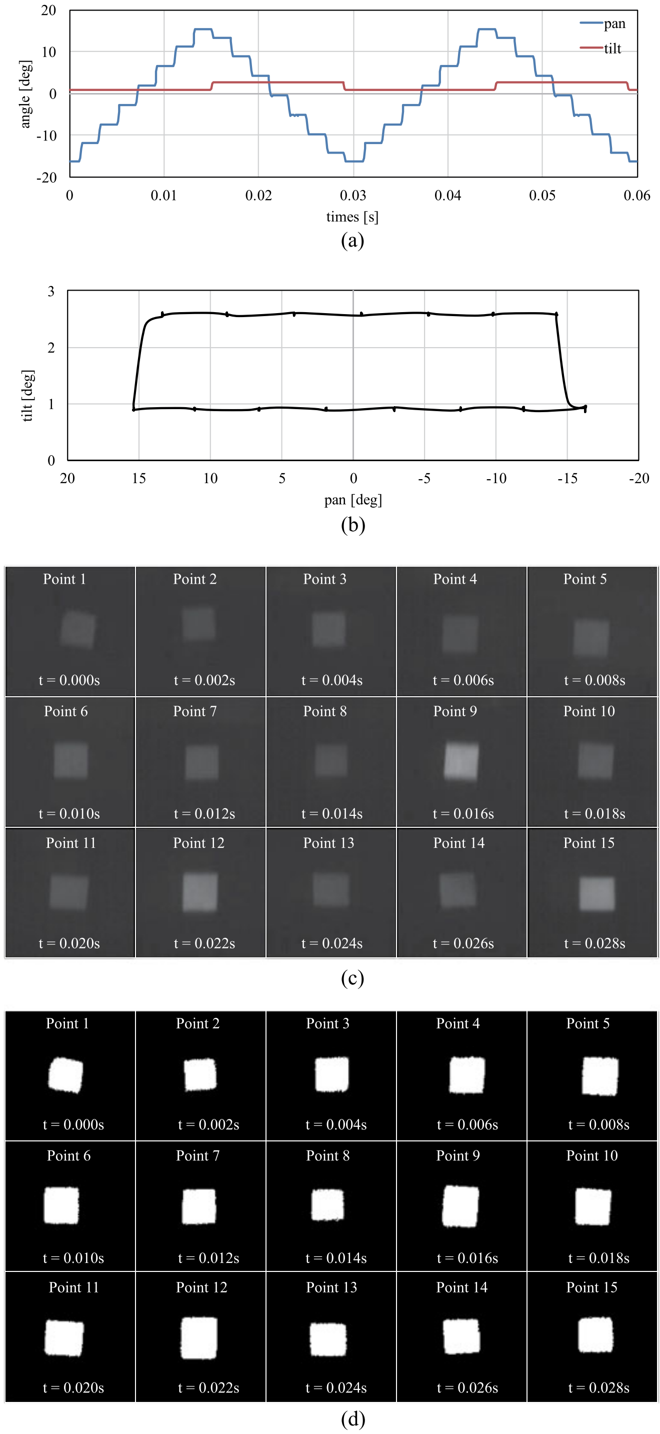


Fig. 3. Pan/tilt angles and captured images in 15-viewpoint switching. (a) Temporal changes. (b) Pan-tilt trajectory. (c) Captured images. (d) Binarized images.

model. Fig. 3(c) and (d) confirms that the images of multiple viewpoints were obtained without motion blur and the images are sufficiently clear to allow binarization. We confirmed that the marker positions calculated from the input images of each viewpoint can be measured within ±0.1° (1 pixel) error when the viewpoints changed every 2 ms for the static bridge model.

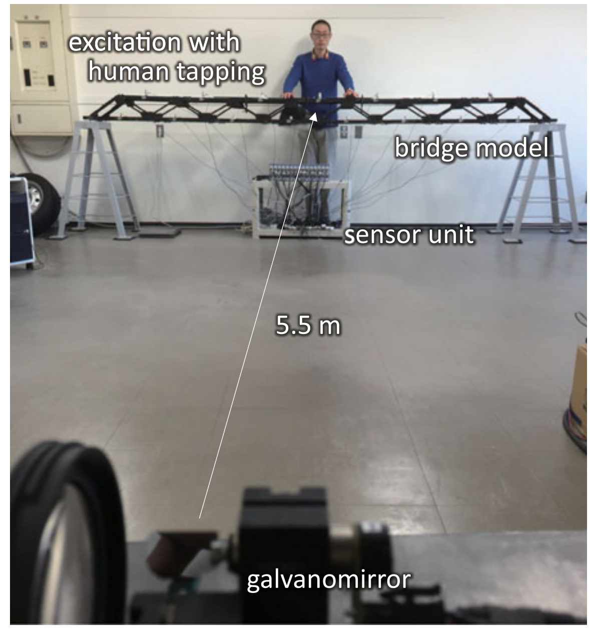


Fig. 4. Experimental scene.

B. Evaluation of the Galvanomirror-Based Multithread Active Vision System

This section describes vibration measurement experiments using the developed galvanomirror-based multithread active vision system, conducted in order to verify whether our system can perform as well as an existing vibration measurement system. The multithread active vision system measured the center of gravity of markers attached at 15 viewpoints in real time. Bridge model vibrations were induced manually as shown in Fig. 4. Fig. 5 shows time series data of the marker's vertical position y of 15 markers calculated from images captured by the high-speed multithread active vision system, with comparative data from the vibration indicator unit. The y position is set to zero when the bridge model is static. The measurement results of the high-speed multithread active vision system are based on data obtained at a sensing rate of 33.3 Hz at each viewpoint. In contrast, the sensing rate of the vibration indicator unit is 2000 Hz. Although the time density of the data from the high-speed multithread active vision is insufficient, it does confirm that the maximum amplitude for the 15 points was 1.5 mm and that damped vibration occurred at around 11 Hz center frequency.

Next, FFT analysis for the sensing data in Fig. 5 was conducted in order to obtain the frequency response. Fig. 6(a) and (b) shows the results of the FFT analysis of the high-speed multithread active vision system and the vibration indicator unit, respectively. A peak frequency of 11.3 Hz was detected from the results of both (a) and (b); this confirms that the peak frequency of the bridge model can be detected from measurement results of the high-speed multithread active vision system just as well as from the data of vibration indicator unit.

Fig. 7 shows time series shape data of the bridge model based on data from the 15 measurement points obtained by the high-speed multithread active vision system. The shape data of the

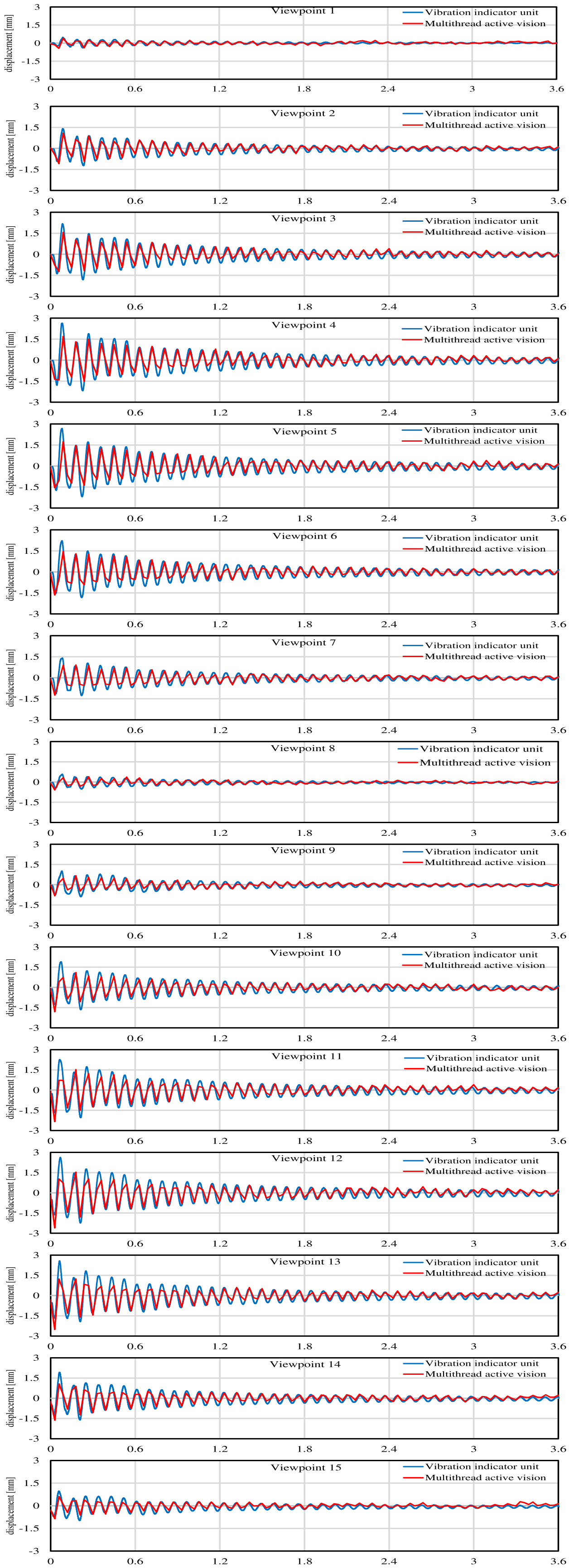


Fig. 5. Vibration displacements at 15 points.

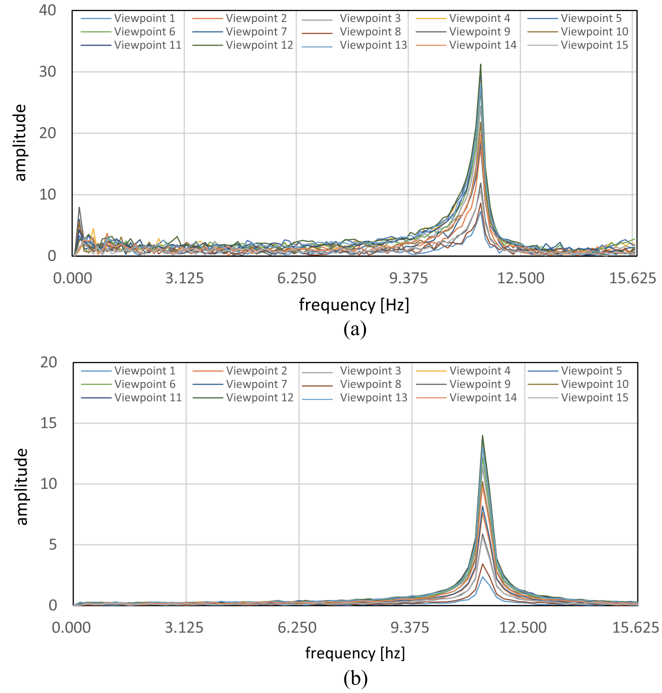


Fig. 6. Frequency response of bridge model. (a) Multithread active vision system. (b) Vibration indicator unit.

bridge model derived from the data of vibration indicator unit is also shown in this figure for comparison. The vertical direction is enlarged 80 times so that vertical vibration displacements are emphasized. Fig. 7(a) shows the shape data without considering synchronization of shooting time for the 15 measurement points in the high-speed multithread active vision sensing system. The shape derived from the high-speed multithread active vision system is completely different from the shape derived from the vibration indicator unit and is physically improbable. In high-speed multithread active vision sensing, the measurement position is sequentially updated every frame interval $\tau = 2$ ms. If the y position of the ith point at time t_k is defined as

$$^{i}Y(t_k) = {}^{i}y(t_k - i\tau) \quad (i = 1, \ldots, 15)$$

30 ms are required to finish sensing the y positions of 15 measurement points. Thus, a synchronization error occurred depending on the sweeping order of the measurement points during the 30-ms interval.

In order to solve this problem, the following third-order spline interpolation with seven interval was used

$$\tilde{y}_i(t) = \mathrm{Spline}(Y_i(t_{k-3}), Y_i(t_{k-2}), Y_i(t_{k-1}), Y_i(t_k)$$
$$Y_i(t_{k+1}), Y_i(t_{k+2}), y_i(Y_{k+3})).$$

Fig. 7(b) is the bridge model shape using $\tilde{y}_i(t)$ in which y was corrected assuming that the vertical positions at the 15 measurement points were measured simultaneously at time t. It is confirmed that the corrected bridge model shape in Fig. 7(b) is the same as the time series shape derived from the vibration indicator unit, without generation of a physically improbable shape.

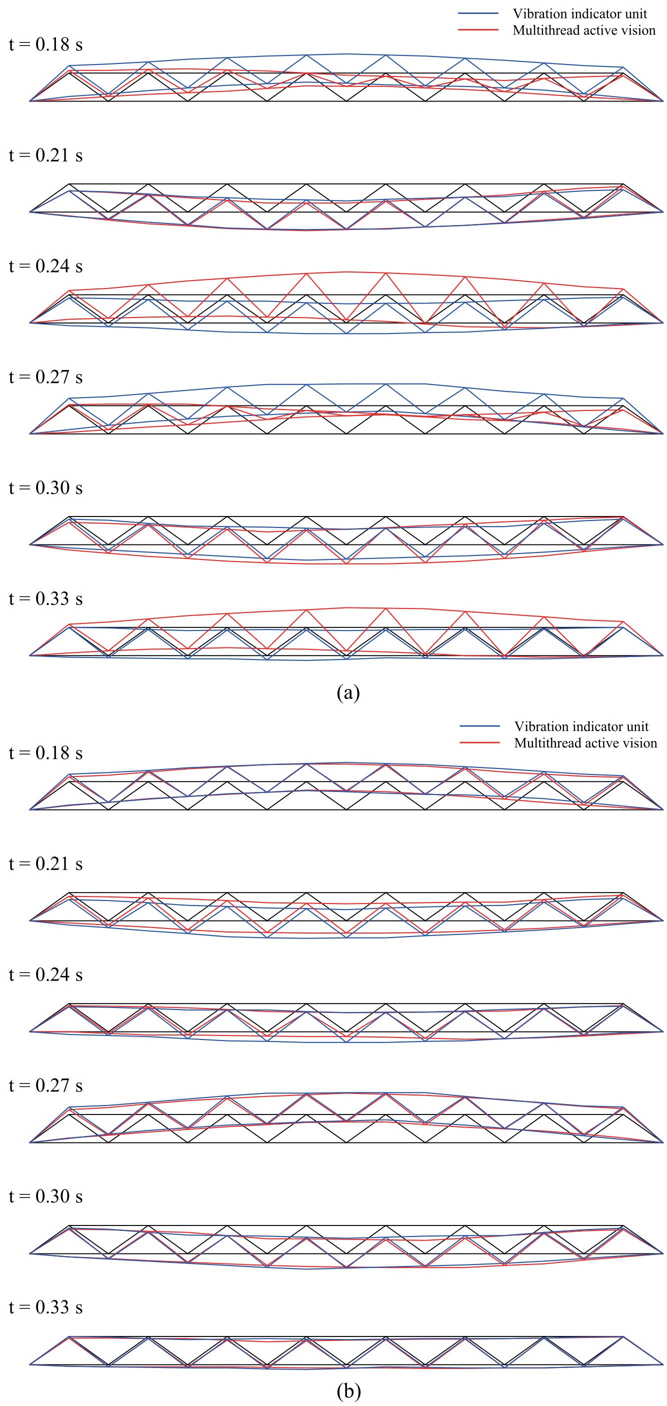


Fig. 7. Spatiotemporal deformation of bridge model. (a) Without calibration. (b) With calibration.

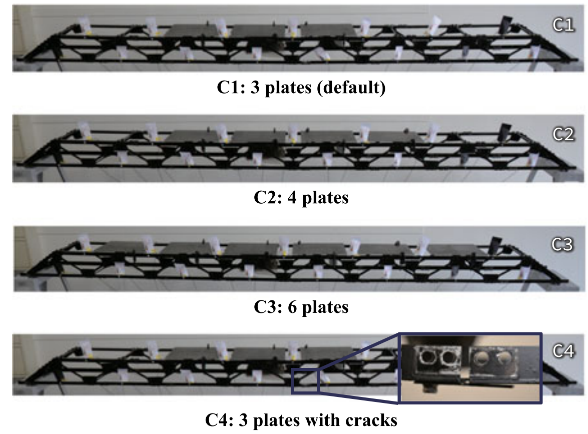


Fig. 8. Conditions of the bridge model.

C. Vision-Based Vibration Sensing Experiment

This section describes vibration measurement experiments for several conditions of the bridge model, conducted in order to verify whether our system can distinguish the condition of bridge via vibration analysis. Fig. 8 shows the bridge model conditions that were applied (C1–C4). In conditions C1–C3, three, four, and six steel plates, respectively, were placed on the center of the bridge model in order to change the weight and resonant frequency of each bridge model. In condition C4, a fraction of the screws were removed so that the bridge model is assumed to be virtually broken.

In these experiments, vibrations for bridge models (C1–C4) were generated manually, as they were in Section IV-B. Fig. 9 shows the spatiotemporal deformation of bridge models (C1–C4) after spline correction, and Fig. 10 shows the first mode shape of bridge models (C1–C4). The modal analysis is conducted based on the SSI-CPAST algorithm [35], [36], which simultaneously identifies the input-invariant dynamic properties of an object under an unknown excitation.

The resonant frequencies of bridge models C1–C3 were determined to be 11.14, 10.42, and 9.65 Hz, respectively (see Table I). It is confirmed that the resonant frequency decreased as the load increased. In contrast, the first mode shapes of bridge models C1–C3 are not significantly different; this means that the dynamic properties of bridge models C1–C3 are not very different. Although the resonant frequencies of bridge models C1–C3 were different owing to the different weights of the bridge models, the mode shapes of the bridge models, which indicate dynamic properties, were almost the same because the bridge models have the same structure. Thus, these sensing results have validity.

The resonant frequency of bridge model C4, in which a fraction of the screws were removed, decreases to a low value compared to the resonant frequency of bridge model C1; also, the first mode shape around the cracked position of bridge model C4 is quite different from the shape of bridge model C1. It is confirmed that the dynamic properties of the bridge model were changed significantly because of the cracks in bridge model C4.

The experimental results in this section are almost same as the analysis results derived from the vibration indicator unit (see Figs. 9 and 10), which confirms the reliability of the analysis results using the proposed multithread active vision system. The measurable range of resonant frequency depends on the virtual PT camera's frame rate in the proposed multithread active vision system. In these experiments, the virtual PT camera's frame rate was 33.3 fps and the maximum resonant frequency was around 11 Hz. Although the time density of the proposed multithread active vision system is lower than that of a vibration indicator unit, the proposed system can measure the displacements

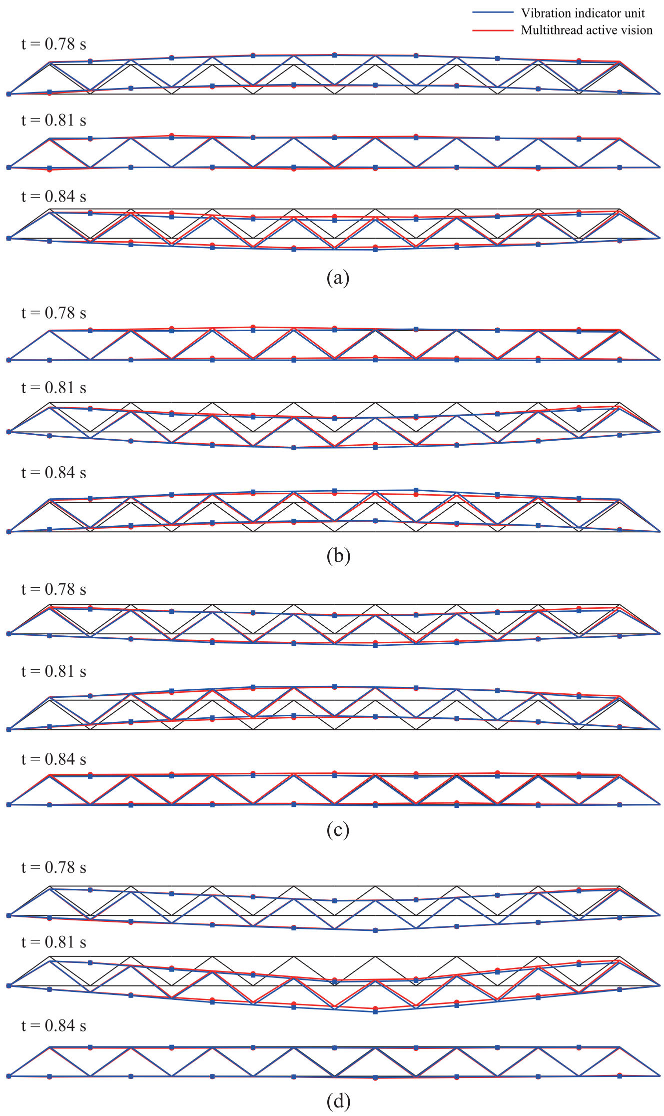


Fig. 9. Spatiotemporal deformation of the bridge models. (a) Deformation of bridge model C1. (b) Deformation of bridge model C2. (c) Deformation of bridge model C3. (d) Deformation of bridge model C4.

of multiple viewpoints using only one camera system and the dynamic properties of the sensing object can be detected from the sensing results of our system with nearly the same accuracy as that of existing multiple sensors provided the resonant frequencies of the sensing objects are low enough compared to the virtual PT camera's frame rate.

V. Consideration for Practical Use

The low brightness, defocus, and accuracy of each measurement point are a matter of concern when we consider applying our proposed sensing method to actual truss bridges dozens of meters long. In this section, we discuss the brightness, defocus, and accuracy of the measurement points. We observed a series of poles set at a distance of 27.5 m from the multithread active vision system and having a width of 20.0 m, as shown in Fig. 11(c), in order to consider the application of our method to actual truss bridges. Corner cubes were used as markers of the measurement points, as shown in Fig. 11(a). Our experiments were conducted using our multithread active vision system with a metal halide lamp and a half-mirror installed to illuminate the markers, as shown in Fig. 11(b). In addition, a zoom lens (JOBLE, 4320–8640 mm, F10-90, HTZ-8600) was attached to the camera head in order to realize long-distance photography.

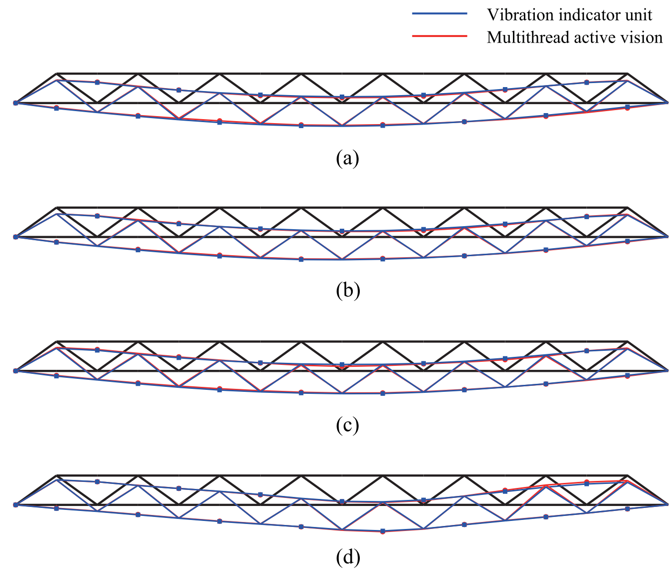


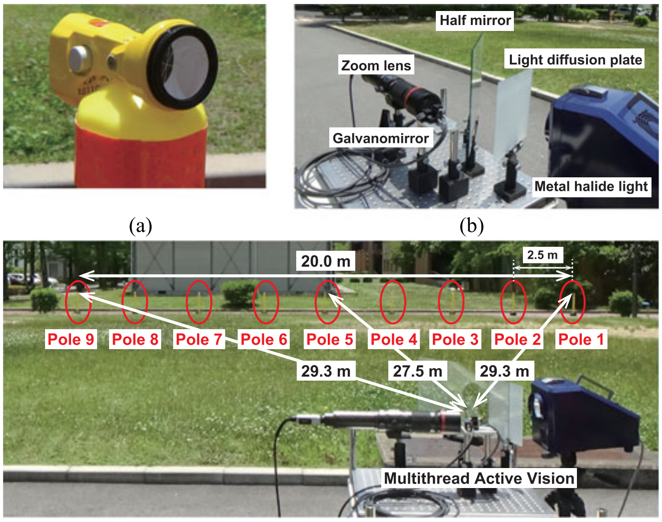
Fig. 10. Mode shape of the bridge models. (a) First mode shape of bridge model C1. (b) First mode shape of bridge model C2. (c) First mode shape of bridge model C3. (d) First mode shape of bridge model C4.

TABLE I
Table of the First Mode Frequency

	C1	C2	C3	C4
Multithread active vision	11.14	10.42	9.65	6.54
Vibration indicator unit	11.25	10.51	9.42	6.54

A. Discussion of Brightness

In long-distance photography, since a low aperture ratio value is in general set when a zoom lens is used, the brightness of the measurement points is insufficient, and thus, it is difficult to apply a vision-based sensing method, including our proposed method. By setting corner cubes or LEDs at the measurement points, observation in order to measure points at a distance of dozens of meters becomes possible. Fig. 12 shows the captured images of markers at a series of nine measurement points installed at a distance of 27.5 m and having a width of 20 m and Fig. 13 shows the binarized images of the nine measurement points. These images confirm that the same algorithm using the proposed system can derive binarized images even on the actual truss-structure scale. We conclude that the brightness is sufficient for measurement when we apply the proposed method to structures having a length of dozens of meters.



(c)

Fig. 11. Experimental scene of supplemental experiments. (a) Corner cube attached on the top of the pole. (b) Overview of the multithread active vision system. (c) Experimental environment.

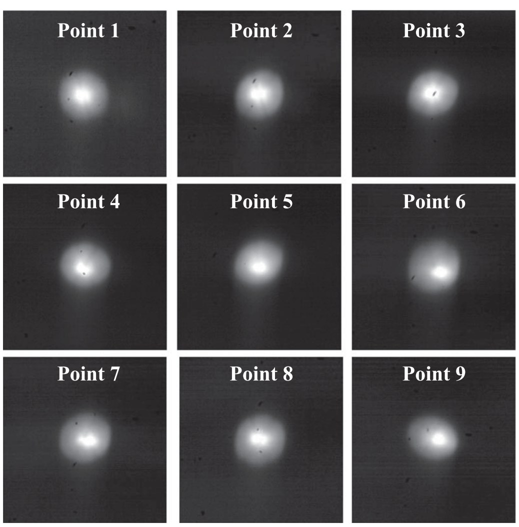


Fig. 12. Captured images.

B. Discussion of Defocus

When structures of a scale of several tens of meters are measured using the proposed method, the distance between the shortest and the longest measuring point from the camera position becomes larger; then, focus blur becomes a matter of concern. When corner cubes or LEDs are set at the measurement points, the proposed method, which derives the position of the measurement point by binarizing the markers at the measurement point, does not cause a serious focus blur problem. The images of the measurement points using corner cubes are binarized correctly, as shown in Fig. 13. The binarized images are not inferior to those in Fig. 3(d) captured from a distance of 5.5 m, and the focus blur is within the allowable range. Thus, we conclude that the defocus is acceptable when applying the proposed method to a structure of a scale of several tens of meters. In addition, we confirmed that it is possible to measure the position of the measurement points, even if they are at a distance of around 40 m.

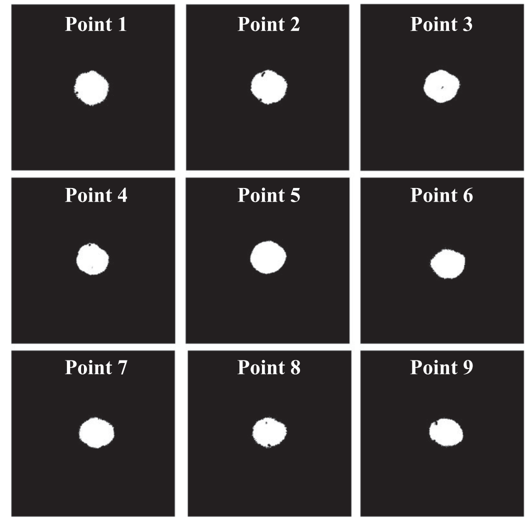


Fig. 13. Binarized images.

C. Discussion of Accuracy

When measuring structures of a scale of several tens of meters using the proposed method, the photographing distance becomes longer and the problem that the measurement accuracy may deteriorate becomes a matter of concern. In this experimental condition, 1 pixel corresponds to 0.04 mm for a 4-m-wide measurement object, whereas 1 pixel corresponds to 0.234 mm for a 20-m-wide measurement object. Although the accuracy deteriorates as the measurement object becomes larger, the amplitude also increases and is more than a few millimeters in the case where the scale of the measurement object is a truss bridge of several tens of meters[39]. Thus, we conclude that the accuracy of the proposed system is sufficient for the measurement of a structure of a scale of several tens of meters.

According to the above three discussions, in our opinion the proposed multithread active vision system can also be applied to the vibration measurement of a truss structure of a scale of several tens of meters.

VI. Conclusion

In this paper, we proposed a concept for a multithread active vision system in which a single high-speed tracking camera can virtually function as multiple PT tracking cameras by accelerating its measurement, computation, and actuation with ultrafast viewpoint switching. We developed a real-time vibration distribution analysis system for wide-range large constructions, which is difficult to measure with one camera system by using a multithread active vision concept that performs virtually as 15 cameras, each operating at 33.3 fps. Vibration measurement experiments were conducted on a 4-m-long truss-structure bridge model to which 15 markers were attached and it was confirmed that the single active vision system could observe the deformation of the bridge structure and estimate modal parameters, such as resonant frequencies and mode shapes, at a frequency of dozens of hertz. Improvement of the sensitivity, sensing accuracy, and sensing range of the proposed high-speed multithread active vision system, and development of applications for system for monitoring society's infrastructure are left for future study.

References

[1] O. S. Salawu and C. Williams, "Review of full-scale dynamic testing of bridge structures," *Eng. Struct.*, vol. 17, no. 2, pp. 113–121, Feb. 1995.

[2] Q. Qin, H. B. Li, L. Z. Qian, and C. K. Lau, "Modal identification of Tsing Ma bridge by using improved eigensystem algorithm," *J. Sound Vib.*, vol. 247, no. 2, pp. 325–341, Oct. 2001.

[3] K. T. Park, S. H. Kim, H. S. Park, and K. W. Lee, "The determination of bridge displacement using measured acceleration," *Eng. Struct.*, vol. 27, pp. 371–378, 2004.

[4] C. Rainieri and G. Fabbrocino, "Automated output-only dynamic identification of civil engineering structures," *Mech. Syst. Signal Process.*, vol. 24, no. 3, pp. 678–695, Apr. 2010.

[5] H. H. Nassif, M. Gindy, and J. Davis, "Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration," *NDT E Int.*, vol. 38, pp. 213–218, 2005.

[6] J. L. Valin, E. Concalves, and J. Davis, "Methodology for analysis of displacement using digital holography," *Opt. Lasers Eng.*, vol. 43, pp. 99–111, 2005.

[7] P. Kohut *et al.*, "Monitoring of a civil structure's state based on noncontact measurements," *Struct. Health Monit.*, vol. 12, pp. 411–429, 2013.

[8] F. Casciati and L. Wu, "Local positioning accuracy of laser sensors for structural health monitoring," *Struct. Control. Health Monit.*, vol. 20, pp. 728–739, 2013.

[9] J. J. Lee and M. Shinozuka, "Real-time displacement measurement of a flexible bridge using digital image processing techniques," *Exp. Mech.*, vol. 46, pp. 105–114, 2006.

[10] D. Feng, M. Q. Feng, E. Ozer, and Y. Fukuda, "A vision-based sensor for noncontact structural displacement measurement," *Sensors*, vol. 15, pp. 16557–16575, 2015.

[11] Y. F. Ji and C. C. Chang, "Nontarget image-based technique for small cable vibration measurement," *J. Bridge Eng.*, vol. 13, pp. 34–42, 2008.

[12] S. W. Kim, B. G. Jeon, N. S. Kim, and J. C. Park, "Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge," *Struct. Health Monit.*, vol. 12, pp. 440–456, 2013.

[13] G. Busca, A. Cigada, P. Mazzoleni, and E. Zappa, "Vibration monitoring of multiple bridge points by means of a unique vision-based measuring system," *Exp. Mech.*, vol. 54, pp. 255–271, 2014.

[14] Y.-Z. Song, C. R. Bowen, A. H. Kim, A. Nassehi, J. Padget, and N. Gathercole, "Virtual visual sensors and their application in structural health monitoring," *Struct. Health Monit.*, vol. 13, pp. 251–264, 2014.

[15] J. J. Lee, H. N. Ho, and J. H. Lee, "A vision-based dynamic rotational angle measurement system for large civil structures," *Sensors*, vol. 12, pp. 7326–7336, 2012.

[16] Y. Fukuda, M. Q. Feng, Y. Narita, S. Kaneko, and T. Tanaka, "Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm," *IEEE Sens. J.*, vol. 13, no. 12, pp. 4725–4732, Dec. 2013.

[17] P. Kohut and P. Kurowski, "Application of modal analysis supported by 3D vision-based measurements," *J. Theor. Appl. Mech.*, vol. 47, no. 4, pp. 855–870, May. 2009.

[18] J. Morlier and G. Michon, "Virtual vibration measurement using KLT motion tracking algorithm," *J. Dyn. Syst., Meas. Control*, vol. 132, no. 1, pp. 011003–011011, Jan. 2010.

[19] P. Avitabile, C. Niezrecki, M. Helfrick, C. Warren, and P. Pingle, "Noncontact measurement techniques for model correlation," *Sound Vib.*, vol. 44, no. 1, pp. 8–12, Jan. 2010.

[20] E. Caetano, S. Silva, and J. Bateira, "A vision system for vibration monitoring of civil engineering structures," *Exp. Tech.*, vol. 35, no. 4, pp. 74–82, Aug. 2011.

[21] J. G. Chen, N. Wadhwa, Y. J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, "Modal identification of simple structures with high-speed vision using motion magnification," *J. Sound Vib.*, vol. 345, pp. 58–71, 2015.

[22] H. Yang, Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "Dynamics-based stereo visual inspection using multidimensional modal analysis," *IEEE Sensors J.*, vol. 13, no. 12, pp. 4831–4843, Dec. 2013.

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

[24] H. Aghajan and A. Cavallaro, Eds., *Multi-Camera Networks: Concepts and Applications*. Amsterdam, The Netherlands: Elsevier, 2009.

[25] X. Wang, "Intelligent multi-camera video surveillance: A review," *Pattern Recognit. Lett.*, vol. 34, pp. 3–19, 2013.

[26] Y. Watanabe, T. Komura, and M. Ishikawa, "955-fps real-time shape measurement of a oving/deforming object using high speed vision for numerous-point analysis," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2007, pp. 3192–3197.

[27] I. Ishii, T. Taniguchi, R. Sukenobe, and K. Yamamoto, "Development of high-speed and real-time vision platform, H3 vision," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 3671–3678.

[28] I. Ishii, T. Taniguchi, K. Yamamoto, and T. Takaki, "High frame-rate optical flow system," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 22, no. 1, pp. 105–112, Jan. 2012.

[29] I. Ishii, T. Ichida, Q. Gu, and T. Takaki, "500-fps face tracking system," *J. Real-Time Image Process.*, vol. 8, pp. 379–388, 2013.

[30] A. Namiki, Y. Imai, and M. Ishikawa, "Development of a high speed multi fingered hand system and its application to catching," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots. Syst.*, 2003, pp. 2666–2671.

[31] T. Aoyama, T. Takaki, T. Miura, Q. Gu, and I. Ishii, "Realization of flower stick rotation using robotic arm," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots. Syst.*, 2015, pp. 5648–5653.

[32] K. Okumura, H. Oku, and M. Ishikawa, "High-speed gaze controller for millisecond-order pan/tilt camera," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 6186–6191.

[33] T. Ueno, Q. Gu, T. Aoyama, T. Takaki, I. Ishii, and T. Kawahara, "Motion-blur-free microscopic video shooting," in *Proc. IEEE Int. Conf. Automat. Sci. Eng.*, 2015, pp. 837–742.

[34] I. Ishii, T. Tatebe, Q. Gu, Y. Moriue, T. Takaki, and K. Tajima, "2000 fps real-time vision system with high-frame-rate video recording," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2010, pp. 1536–1541.

[35] H. Yang, T. Takaki, and I. Ishii, "Simultaneous dynamics-based visual inspection using modal parameter estimation," *J. Robot. Mechatronics*, vol. 23, no. 1, pp. 180–195, 2011.

[36] H. Yang, Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "Dynamics-Based stereo visual inspection using multidimensional modal analysis," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4831–4843, 2013.

[37] J. P. Lynch and K. J. Loh, "A summary review of wireless sensors and sensor networks for structural health monitoring," *Shock Vib. Dig.*, vol. 38, no. 2, pp. 91–128, 2006.

[38] S. Cho, C. Yun, J. P. Lynch, A. T. Zimmerman, B. F. Spencer, Jr, and T. Nagayama, "Smart wireless sensor technology for structural health monitoring of civil structures," *Steel Struct.*, vol. 8, pp. 267–275, 2008.

[39] Sh. Ataei, A. A. Aghakouchak, M. S. Marefat, and S. Mohammadzadeh, "Sensor fusion of a railway bridge load test using neural networks," *Expert Syst. Appl.*, vol. 29, pp. 678–683, 2005.

Tadayoshi Aoyama (M'12) received the B.E. degree in mechanical engineering, the M.E. degree in mechanical science eingineering, and the Ph.D. degree in micro-nano systems engineering from Nagoya University, Nagoya, Japan, in 2007, 2009, and 2012, respectively.

He is currently an Assistant Professor with the Department of Micro-Nano Mechanical Science and Engineering, Nagoya University. From 2012 to 2017, he was an Assistant Professor at Hiroshima University, Higashi-Hiroshima, Japan. His research interests include vision-based sensing, dynamic motion control, and human-assistive robots.

Liang Li received the B.S. degree in electronic information science and technology from China Agricultural University, Beijing, China, in 2006, and the M.E. degree in system cybernetics from Hiroshima University, Higashi-Hiroshima, Japan, in 2017.

From 2009 to 2013, he was in the Beijing Luster Light Tech Group, Beijing. Since 2017, he has been with Hitachi High-Technologies Corporation, Tokyo, Japan. His research interest includes vision-based modal analysis, image processing, and recognition.

Mingjun Jiang received the B.E. degree in computer science and technology from the Hefei University of Technology, Hefei, China, in 2010; the M.E. degree in software engineering from Harbin Engineering University, Harbin, China, in 2014; and the Ph.D. degree in system cybernetics from Hiroshima University, Higashi-Hiroshima, Japan, in 2017.

Since 2017, he has been a Designated Assistant Professor with the Department of System Cybernetics, Hiroshima University. His research interests include high-speed vision and vibration source localization.

Kazuki Inoue received the B.E. degree in system engineering from Hiroshima University, Higashi-Hiroshima, Japan, in 2016, where he is currently working toward the M.E. degree in system cybernetics.

His current research focuses on sensory information processing.

Takeshi Takaki (M'06) received the B.E. and M.E. degrees in mechanical engineering from the Tokyo University of Science, Tokyo, Japan, in 2000 and 2002, respectively, and the Ph.D. degree in mechano-micro engineering from the Tokyo Institute of Technology, Tokyo, in 2006.

Since 2011, he has been an Associate Professor with Hiroshima University, Higashi-Hiroshima, Japan. From 2002 to 2003, he was at Disco Crop., Tokyo. From 2006 to 2007, he was a Researcher at the Tokyo Institute of Technology. From 2007 to 2011, he was an Assistant Professor at Hiroshima University. His research interests include mechanical design, robotic arm, and sensory information processing.

Idaku Ishii (M'06) received the B.E., M.E., and Ph.D. degrees in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1992, 1994, and 2000, respectively.

He is currently a Professor with Hiroshima University, Higashi-Hiroshima, Japan. From 1996 to 2000, he was a Research Associate at the University of Tokyo. In 2000, he joined the Tokyo University of Agriculture and Technology, Tokyo, as a Lecturer, and from 2002 to 2003, he was an Associate Professor. From 2003 to 2007, he was an Associate Professor at Hiroshima University. His research interests include high-speed vision, sensory information processing, and its applications.

Hua Yang (M'13) received the B.E. degree in mechanical design-manufacture and its automation, and the M.E. degree in mechatronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2008, respectively, and the Ph.D. degree in system cybernetics from Hiroshima University, Higashi-Hiroshima, Japan, in 2011.

From 2011 to 2012, he was at Hiroshima University as a Research Associate, and was a special-appointment Assistant Professor in April 2012. He is currently an Associate Professor with the School of Mechanical Science and Engineering, Huazhong University of Science and Technology. His current research interests include object recognition, detection and tracking, particle image velocity, and dynamics-based vision inspection.

Chikako Umemoto received the B.E degree in civil engineering from Hiroshima University, Higashi-Hiroshima, Japan, in 1995.

In 1995, she joined Keisoku Research Consultant Corporation, Hiroshima, Japan. Her research interest includes maintenance and monitoring engineering of civil structures.

Hiroshi Matsuda received the B.E. and M.E. degrees from the Kyushu Institute of Technology, Kitakyushu, Japan, in 1980 and 1982, respectively, and the Ph.D. degree from Kyushu University, Fukuoka, Japan, in 1987, all in civil engineering.

He is currently a Professor with Nagasaki University, Nagasaki, Japan, where he was a Research Associate from 1982 to 1988, a Lecturer from 1988 to 1991, and an Associate Professor from 1991 to 2003. His research interests include structural mechanics, bridge engineering, and maintenance and management engineering.

Makoto Chikaraishi received the B.E. degree in civil and environmental engineering, and the M.E. and Ph.D. degrees in international development and cooperation from Hiroshima University, Higashi-Hiroshima, Japan, in 2005, 2007, and 2010, respectively.

He is currently an Associate Professor with Hiroshima University. From 2010 to 2012, he was an Assistant Professor at Hiroshima University. His research interests include urban and transportation issues, and activity-travel behavior modeling with applications in various fields.

Akimasa Fujiwara received the B.E. degree in civil engineering, and the M.E. and Ph.D. degrees in civil and environmental engineering from Hiroshima University, Higashi-Hiroshima, Japan, in 1983, 1985, and 1993, respectively.

He is currently a Professor with Hiroshima University. From 1985 to 1993, he was a Research Associate at the Kure National College of Technology, Kure, Japan. In 1993, he joined Hiroshima University as a Research Associate, and he was an Associate Professor from 1994 to 2002. His research interests include transportation planning, transportation policy evaluation, and sustainable development and transportation.