Self-Position Measurement of Multi-Copter for Photographing under the Bridge Girder

Tetsuya Morisaki [1*], National Institute of Technology, Ube college, Japan

Hisao Emoto [2], National Institute of Technology, Fukushima college, Japan

Yutaka Watanabe [3], Luce Search Co., Ltd., Japan

Abstract: Much attention has been focused on the use of a camera mounted on a multi-copter for visual inspection. However, in the case of photographing under a bridge girder, the operation of the multi-copter is required a skillful technique because the operator controls it while being careful about avoiding contacts with the bridge. Furthermore, the global positioning system (GPS) cannot be used because the wave is often jammed by the bridge girder. For assisting the operator, we propose the application for the perspective three-point method as the self-position measurement method. The measurement system consists of a landmark set on the side of a bridge girder and a pan-tilt camera which is mounted on the multi-copter for photographing the landmark (apart from the one for photographing the bridge). In order to present the guideline for estimating the position error of the system, we analyzed the effect of an image processing error on the accuracy of the calculated self-position of the pan-tilt camera. Based on the result, we employed a three-dimensional regular tetrahedron landmark on the system and evaluated it in experiments. The experimental results show that the proposed system is effective and practical for the operation support under conditions in which GPS cannot be used.

Key-Words: *Multi-copter, Bridge Girder, Bridge Maintenance, Visual Inspection, Perspective-Three-Point*

1. Introduction

Experts perform visual inspection of general bridges by approaching and analyzing them or by analyzing changes in them, such as cracks and lifts, in photographs and videos so that the relevant data can be linked to maintenance management [1]. In order to detect and evaluate the anomalous condition, digital image processing methods have been developed. Schutter developed and demonstrated a video microscope crack monitoring system [2]. Yamaguchi and Hashimoto introduced an efficient and high-speed crack detection method using percolation-based image processing [3]. Adhikari et al. proposed an integrated model based on digital image processing which consists of crack quantification, change detection, neural networks, and 3D visualization models [4]. Valenca et al. proposed a crack characterizing method using photogrammetry and image processing [5].

However, as shown in Fig.1, when inspecting the lower surface of bridge girders, the challenge is introducing time- and labor-saving technologies for securing the safety of experts who perform visual inspection and of workers in charge of photographing. To overcome such problems, promising approaches have been proposed. Oh et al. developed a robotic system for bridge inspection for practical use with automatic and manual (tele-operation) inspection mode [6]. The system writes the bridge safety diagnosis reports from the image capturing. Li et al. presented a

Received: 2019/12/10, Accepted: 2019/12/27
*Corresponding author: Tetsuya Morisaki
Email: morisaki@ube-k.ac.jp

Figure 1. Photographing by a labor on a temporary scaffold deck for a visual inspection.

Figure 2. Photographing for an inspection of a lower surface of a bridge girder with manual multi-copter. (a) A crack photo from the multi-copter. (b)The actual situation of operator.

long-distance precision inspection method with images obtained by a telephoto lens and CCD camera [7]. By limiting steel bridges, robots with magnetic wheels can be utilized [8-11].

In recent years, much attention has been focused on the use of multi-copters for visual inspection, mainly because they can start vertical and horizontal flight from the hovering state. The mere installation of a camera on these multi-copters without a foothold enables close-up inspection. Figure 2(a) shows the example of a photograph for an inspection of a lower surface of a bridge girder with manual multi-copter flights in the past and Figure 2(b) shows the operator. To position a multi-copter close to the lower surface of a bridge girder, the operator must enter the river. In this case, the operator must be skilled at maneuvering the multi-copter. In addition, ensuring the operator's safety is important. To solve this problem, it is necessary to support the operation of multi-copters for acquiring the position information and performing automatic navigation. The global positioning system (GPS) is used to acquire the position information of many multi-copters. However, the radio waves employed by this system deteriorate markedly under the lower surface of a bridge girder, making it difficult to acquire accurate position information. As a substitute for GPS, the autonomous flight of a multi-copter can be controlled by either photographing a landmark mounted on the multi-copter using a camera suspended from the bridge [12] or by extracting feature points from the environmental images captured using a camera mounted on the multi-copter [13]. In the method of recognizing feature points in such images and measuring their positional relationship using a camera, the accuracy of the feature points extracted from the images influences the positional accuracy and the magnitude of this influence is related to the locations of the feature points. In this study, as an approach to arrange optimal feature points, we proposed a self-position measurement system that uses landmark with feature points and a camera mounted on a multi-copter flying under the lower surface of a bridge girder to estimate the impact of errors that occur during the extraction of feature points from the landmark image on the self-position measurement of multi-copters. Based on our findings, we present guidelines for designing the landmark.

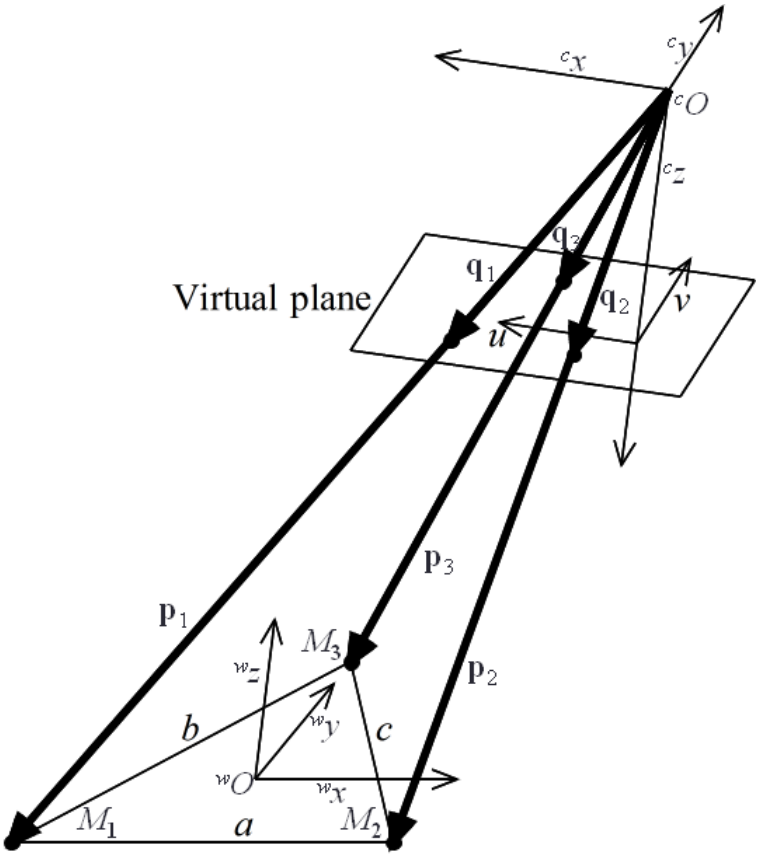

Figure 3. Perspective-Three-Point and definition of coordinate systems.

2. Position measurement performed using the perspective three-point method

Given the limited time available for bridge inspection, it should be possible to install and uninstall the inspection system within a short amount of time. Therefore, in this paper, we employed a landmark that can be simply placed by hanging it from the bridge. The multi-copter is equipped with a camera (apart from the one for photographing the bridge) for capturing pictures of the landmark and a pan–tilt mechanism so that the camera can recognize the landmark while flying in a wide three-dimensional (3D) space. The several feature points on the landmark help the multi-copter to identify its own position from the photographed images of the feature points. Three or more feature points are necessary assuming that the multi-copter may perform rotational motion as well. Therefore, we tested the proposed method using three feature points because the smaller the number of feature points, the shorter the image processing time for recognizing and extracting feature points from the landmark image.

The perspective three-point (P3P) method is a technique of identifying the position of a camera from the images obtained by photographing three feature points that share a known positional relationship in the landmark [14, 15]. Figure 3 outlines the P3P method. M_i (i = 1-3) represents three feature points on the landmark, u and v represent the image coordinate system, and M_i (i = 1-3) is projected on the virtual plane by perspective transformation. ($^c x$, $^c y$, $^c z$) and $^c O$ are the camera coordinate system, and its origin is defined such that u and $^c x$ as well as v and $^c y$ are in parallel, respectively, and the origin of the image coordinate system is on $^c z$. Here, if the distance between $^c O$ and the origin of the image coordinate system is 1, $\mathbf{q}_i$ (i = 1–3) is expressed as follows:

$$\mathbf{q}_i = \begin{bmatrix} u_i \cdot u_s \\ v_i \cdot v_s \\ 1 \end{bmatrix} \quad (i = 1\text{–}3), \tag{1}$$

where (u_i, v_i) (i = 1–3) is a coordinate on the image of M_i (i = 1–3) and (u_s, v_s) is a coefficient obtained by camera calibration. Furthermore, if s_i (i = 1–3) is assumed to be an arbitrary real number, we have the following relationship:

$$\mathbf{p}_i = s_i \mathbf{q}_i . \quad (i = 1\text{–}3) \tag{2}$$

Let a, b, and c be the length of each side of the triangle comprising M_i (i = 1–3). Then, we have:

$$\begin{cases} \left| s_1\mathbf{q}_1 - s_2\mathbf{q}_2 \right|^2 = a^2 \\ \left| s_1\mathbf{q}_1 - s_3\mathbf{q}_3 \right|^2 = b^2 \\ \left| s_2\mathbf{q}_2 - s_3\mathbf{q}_3 \right|^2 = c^2 \end{cases} \tag{3}$$

and $\mathbf{p}_i$ (i = 1–3) in Eq. (2) is obtained by solving the simultaneous equations of s_i (i = 1–3). In the calculation discussed below, s_i (i = 1–3) is obtained numerically using the Newton–Raphson method and the termination condition of iterative calculation is set using the update amount Δs_i (i = 1–3) of s_i (i = 1–3), which is assumed to satisfy the following equation:

$$\sum_{i=1}^{3} \Delta s_i^{\,2} < 10^{-5} . \tag{4}$$

Then, we describe the procedure of obtaining the coordinates of ${}^{c}O$ in the real coordinate system or the self-position information wrc of the camera from $\mathbf{p}_i$ (i = 1–3). Here, we assume that the origin of the real coordinate system ${}^{w}O$ is the centroid of M_i (i = 1–3) and that each coordinate axis (^{w}x, ^{w}y, ^{w}z) is defined by assuming that ^{w}x is in the direction from M1 to M2, ^{w}y is in the direction of M3, and ^{w}z is directed vertically with respect to ^{w}x and ^{w}y. When the origin of the real coordinate system ${}^{w}O$ is expressed as ${}^{c}\mathbf{r}_w$ in the camera coordinate system, it forms the centroid of M_i (i = 1–3) and the relationship holds:

$${}^{c}\mathbf{r}_w = \frac{\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3}{3} \tag{5}$$

By expressing the unit vectors along the ^{w}x, ^{w}y, and ^{w}z axes in the camera coordinate system as i, j, and k , respectively, we have:

$$\mathbf{i} = \frac{1}{\left| \mathbf{p}_2 - \mathbf{p}_1 \right|} \left(\mathbf{p}_2 - \mathbf{p}_1 \right), \tag{6}$$

$$\mathbf{j} = \frac{1}{\left| \mathbf{p}_3 - {}^{c}\mathbf{r}_w \right|} \left(\mathbf{p}_3 - {}^{c}\mathbf{r}_w \right), \tag{7}$$

$$\mathbf{k} = \mathbf{i} \times \mathbf{j} . \tag{8}$$

Using these vectors, ${}^{w}\mathbf{r}_c$ can be expressed as follows:

$${}^{w}\mathbf{r}_c = \begin{pmatrix} \mathbf{i} \cdot {}^{c}\mathbf{r}_w \\ \mathbf{j} \cdot {}^{c}\mathbf{r}_w \\ \mathbf{k} \cdot {}^{c}\mathbf{r}_w \end{pmatrix} . \tag{9}$$

3. Influence of errors occurring during feature point extraction

The (u_i, v_i) (i = 1–3) are the camera coordinates of the feature points extracted from the image of a landmark, and they inevitably contain the errors that occur during image processing. Therefore, to clarify the influence of errors on self-position measurement using a multi-copter, a simulation was performed, as described in the following section, by assuming an equilateral triangle with side l as the landmark.

First, along with the movement of a multi-copter, we rotate the landmark assuming a change in the viewing angle of the mounted camera that follows the landmark using the pan–tilt mechanism. Then, the coordinates (u_i, v_i) (i = 1–3) of the three vertices on the image are calculated based on the positional relationship between the origin of the real coordinate system ${}^{w}O$, origin of the camera coordinate system ${}^{c}O$, and rotation of each axis. By applying (u_i, v_i) (i = 1–3) to the aforementioned P3P method, we find the vector ${}^{w}\mathbf{r}_c$ from the origin of the real coordinate system ${}^{w}O$ in the real coordinate system to the origin of the camera coordinate system ${}^{c}O$. Then, assuming that an error, (e_{ui}, e_{vi}) (i = 1–3), occurs during the extraction of the feature points, the error term is added to (u_i, v_i) (i = 1–3). In the same manner, using

the P3P method, we obtain the vector ${}^{w}\mathbf{e}_c$ from the origin of the system ${}^{w}O$ to the origin of the camera coordinate system ${}^{c}O$. The error r in the self-position measurement caused due to the extracted (e_{ui}, e_{vi}) (i = 1–3) is defined as follows:

$$r = \left| {}^{w}\mathbf{r}_c - {}^{w}\mathbf{e}_c \right|. \quad (10)$$

In the Newton–Raphson method, the initial value of s_i (i = 1 to 3) was considered as the true value obtained from the positional relationship between the origin of the real coordinate system ${}^{w}O$, origin of the camera coordinate system ${}^{c}O$, and rotation of each axis. Considering the portability of the landmark, the size of the equilateral triangle with vertices M_i (i = 1–3) was determined by assuming the length of one side (l = 1000 mm). Furthermore, we assumed (u_s, v_s) = (0.500 × 10^{-3} mm/pixel, 0.500 × 10^{-3} mm/pixel) for a commercial web camera. To examine the influence of the change in the viewing angle caused due to the pan–tilt function, we began working with the movement of the multi-copter and the camera following the landmark. The center of gravity of the equilateral triangle whose base is a side comprising M_3 and M_1 was placed at the center of the image, and M_i (i = 1–3) was projected as an equilateral triangle. Then, a rotation of θ° was assumed around the axis formed by translating the x axis ${}^{c}x$ of the camera coordinate system by 10,000 mm against the origin of the real coordinate system ${}^{w}O$. Similarly, a rotation of ϕ° was assumed around the axis formed by translating the y axis ${}^{c}y$. Finally, the measurement error r was calculated. While extracting the feature point, we assumed that the error (e_{ui}, e_{vi}) (i = 1–3) was within ±5 pixels with a probability of 95% and used normally distributed random numbers with a standard deviation of 2.551 pixels.

Under the aforementioned conditions, the process of calculating the measurement error r was repeated 2^{15} (= 32768) times. Table 1 lists the average (E) and variance (V) values obtained at (θ, ϕ) = (−60, 0), (−30, 0), (−15, 0), (−1, 0, 0), and (0, 0). Figure 4 shows the probability density distribution of r. A comparison with the gamma distribution γ (r: α, β), which was prepared using the shape parameter α (= E/β) and the scale parameter β (= V/E) extracted from Table 1, shows that the distribution of the self-position measurement error closely resembles the gamma distribution.

Table 1 The result of average of the estimating error.

θ°	−60	−30	−15	0
E mm	345	634	1225	2271
V mm^2	23924	100438	378991	406725

Then, to examine the behavior of the errors, including ϕ, the calculation was performed over a wide range of θ and φ (−85°–85°) with a step of 1°. We calculated r 2^{10} (= 1024) times for each value of θ and ϕ, and Fig. 5 shows the average (E) value of r. This result shows that as both θ and ϕ approach 0 or as the camera and the equilateral triangle come closer to each other, the measurement error r increases, indicating that it is desirable to install the landmark to avoid such a scenario.

4. Consideration of 3D landmark

Although the number of extracted feature points increases by one, using a regular tetrahedron landmark, an equilateral triangle suitable for self-position measurement wherein the camera and the equilateral triangle do not face each other can be selected from the four equilateral triangles constituting a regular tetrahedron. To evaluate this tradeoff, the measurement error was calculated for a case wherein the landmark is a three-dimensional regular tetrahedron by adding a feature point, M_4, on the ${}^{w}z$ axis. From the four equilateral triangles constituting a regular tetrahedron, the triangle to be used in the P3P method was selected as follows. First, (u_4, v_4) was extracted for M_4 in a way same as that considered for extracting (u_i, v_i) (i = 1–3). Then, the areas S_i (i = 1–4) pixel^2 are calculated for the following triangle T_i (i = 1–4):

$T_1\{(u_1+e_{u1},v_1+e_{v1}),(u_2+e_{u2},v_2+e_{v2}),(u_3+e_{u3},v_3+e_{v3})\}$,

$T_2\{(u_1+e_{u1},v_1+e_{v1}),(u_2+e_{u2},v_2+e_{v2}),(u_4+e_{u4},v_3+e_{v4})\}$,

$T_3\{(u_1+e_{u1},v_1+e_{v1}),(u_3+e_{u3},v_3+e_{v3}),(u_4+e_{u4},v_3+e_{v4})\}$,

$T_4\{(u_3+e_{u3},v_3+e_{v3}),(u_2+e_{u2},v_2+e_{v2}),(u_4+e_{u4},v_3+e_{v4})\}$.

Here, (e_{u4}, e_{v4}) is a random number generated in a manner similar to that considered for generating (e_{ui}, e_{vi}) (i =

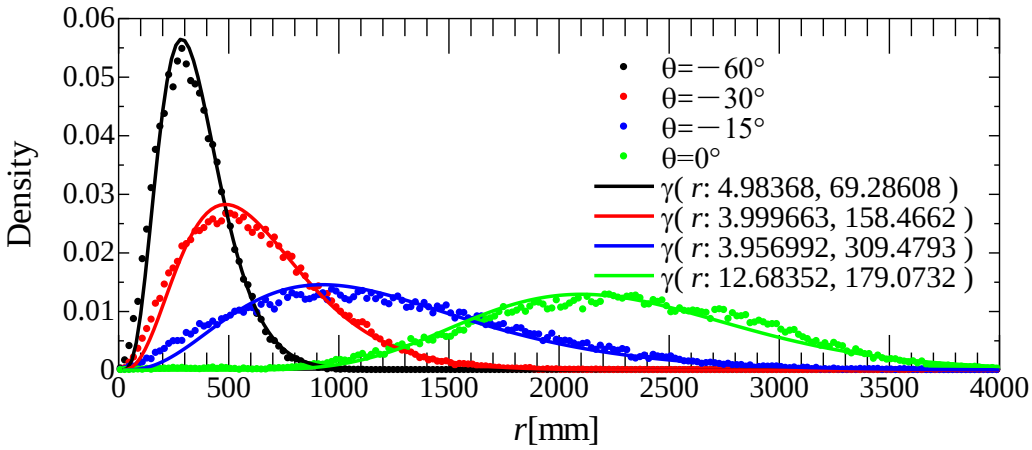

Figure 4. The approximation of error histogram obtained by P3P to gamma distribution.

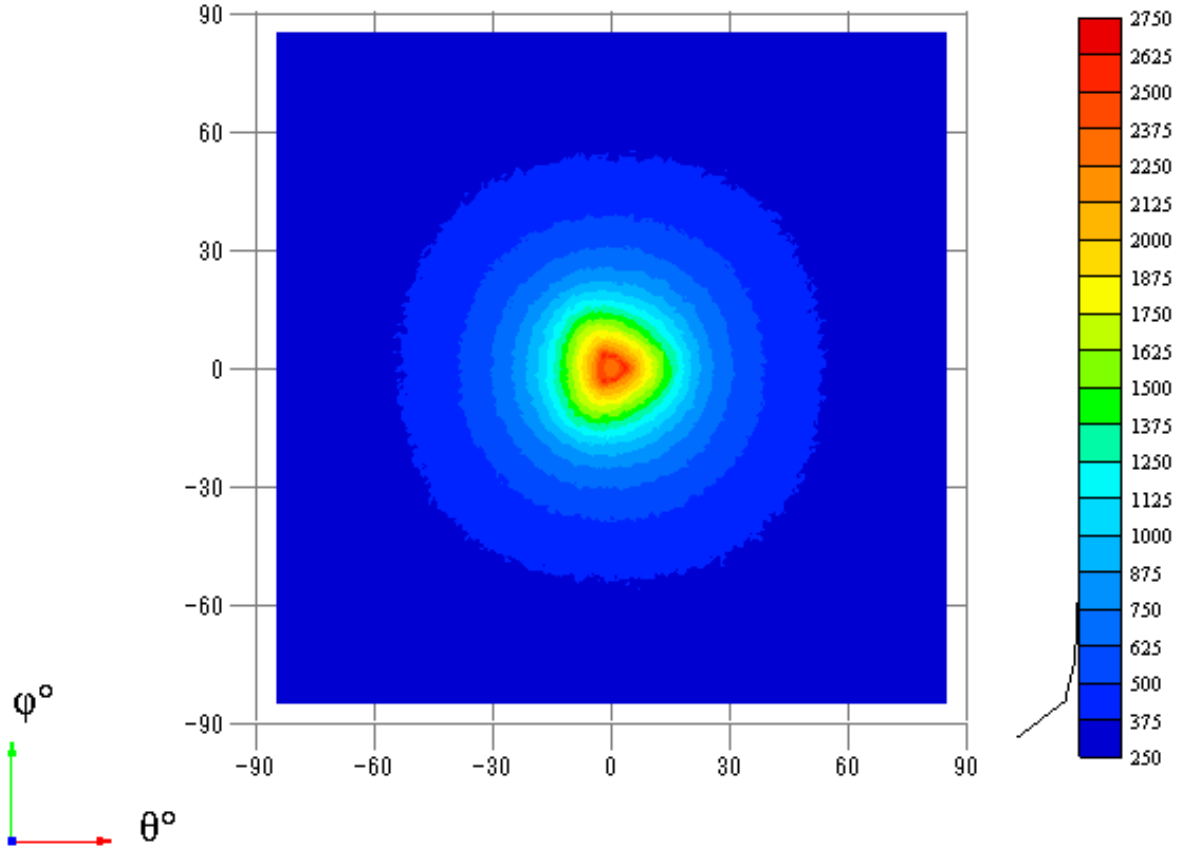

Figure 5. The result of the average of the estimating error E[mm] obtained by P3P assuming the regular triangle landmark.

1–3). The minimum value of S_{min} was obtained from the obtained S_i (i = 1–4) values, and three points constituting the triangle T_{min} were applied to the P3P method. Similar to the case of Fig. 5, the process of calculating r 2^{10} (= 1024) times was repeated and the average (E) value is shown in Fig. 6(a). E peaked at (θ, ϕ) = (57°, −3°), which represents a great improvement, so we modified the graph legends to make it easier to understand the distribution behavior. In addition, in Fig. 6(b), the standard deviation σ (mm) at the same instance is shown so it can be used as an index to estimate the

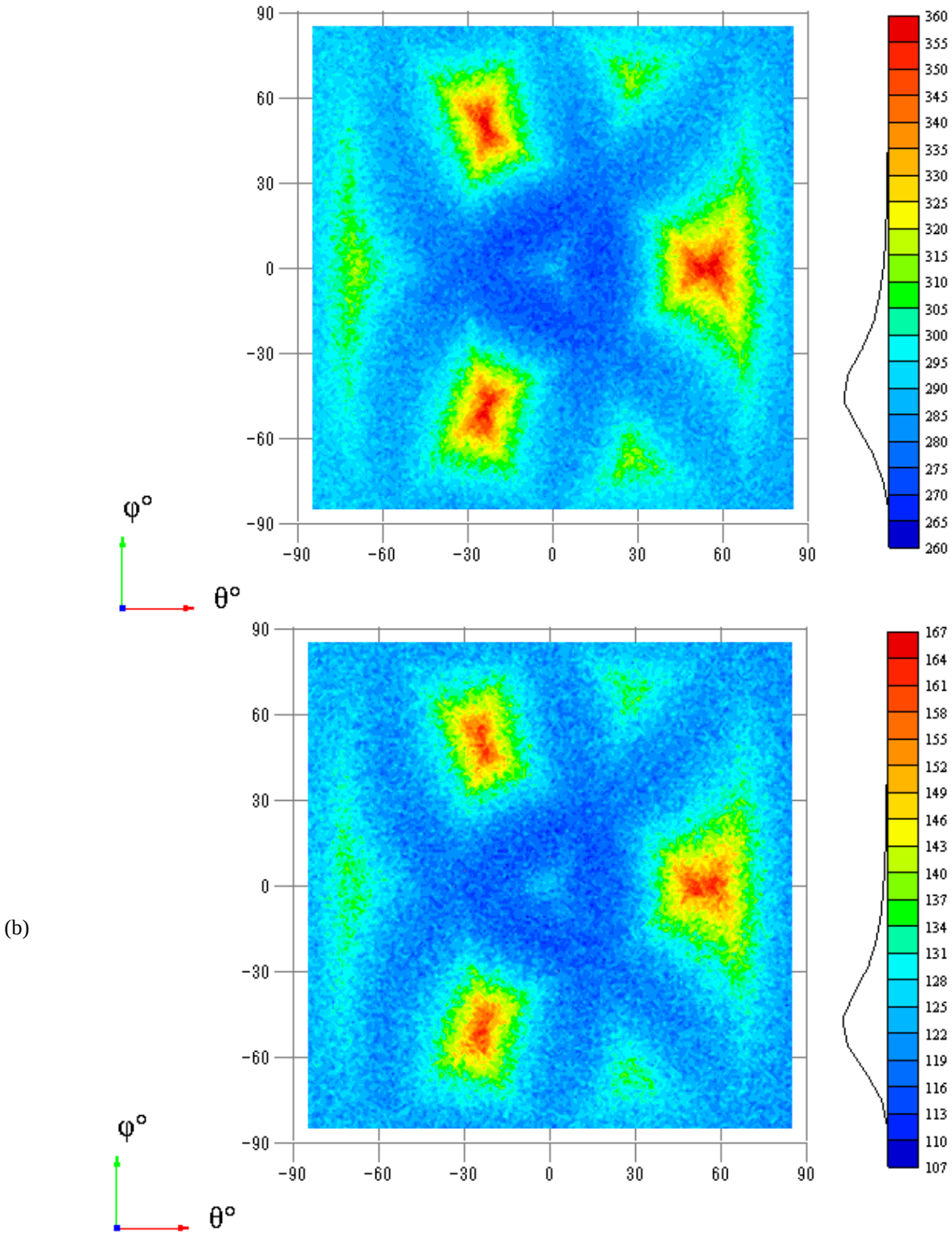

Figure 6. Results of selecting the triangle from the regular tetrahedron landmark for P3P solution. (a) Distribution of average of estimating error E mm. (b) Distribution of standard deviation σ mm.

measurement error. The self-position measurement error can be estimated based on the obtained results. At (θ, ϕ) = (57°, −3°), where the average (E) value peaked when a regular tetrahedral landmark was used, we estimated that r falls within the range of 674.1 mm with 95% probability considering the average (E) value (= 364.0 mm) and standard deviation σ (= 166.6 mm) shown in Fig. 6 and gamma distribution.

For the verification of the proposing method, we experimented with the regular tetrahedral landmark model with a side of 400mm shown in Fig.7. The landmark consists of three red balls and one green ball with a diameter of 40 mm. Fig.8 shows the example of an image taken with the pan–tilt camera (QCAM-200RX, Logicool) set on ^{w}z=1350mm used in the experiment and the result of an image processing for detecting feature points. Figure 9 shows 15 positions where the pan–tilt camera was set and measuring results from images of the landmark taken 90 times in each position. When we collected photographs for an inspection of a lower surface of a bridge girder with multi-copter flights in the past as shown in Fig.2, the multi-copter was operated not to approach within 3m to the bridge girder for avoiding any accidental fall.

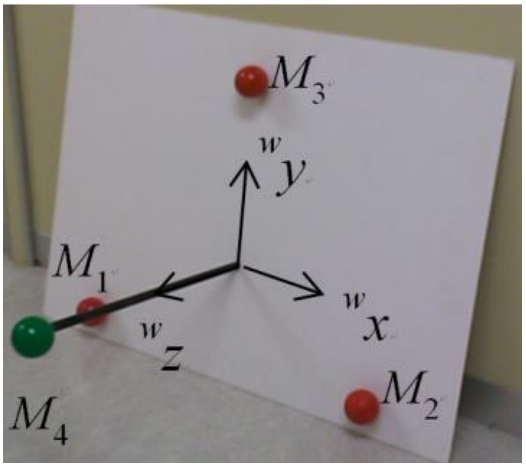

Figure 7. The regular tetrahedral landmark used in the experiment.

(a)

(b)

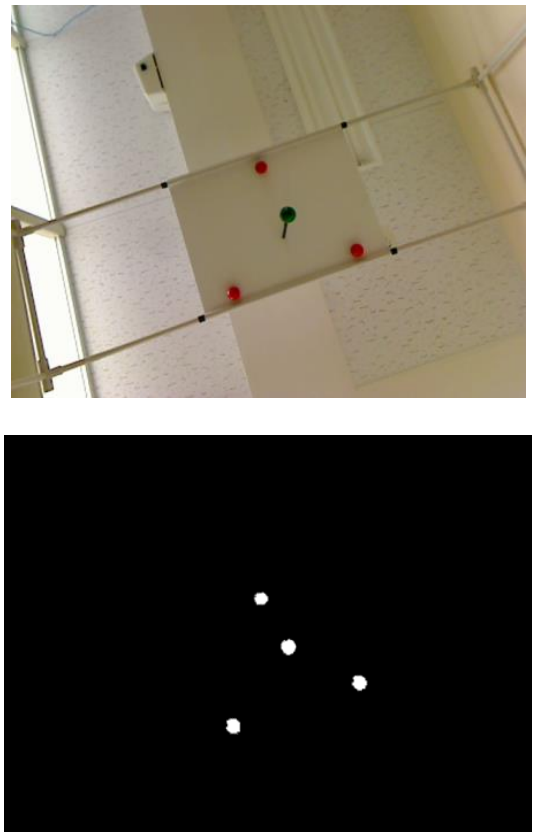

Figure 8. The example of the image processing. (a) Source image. (b) detected feature points.

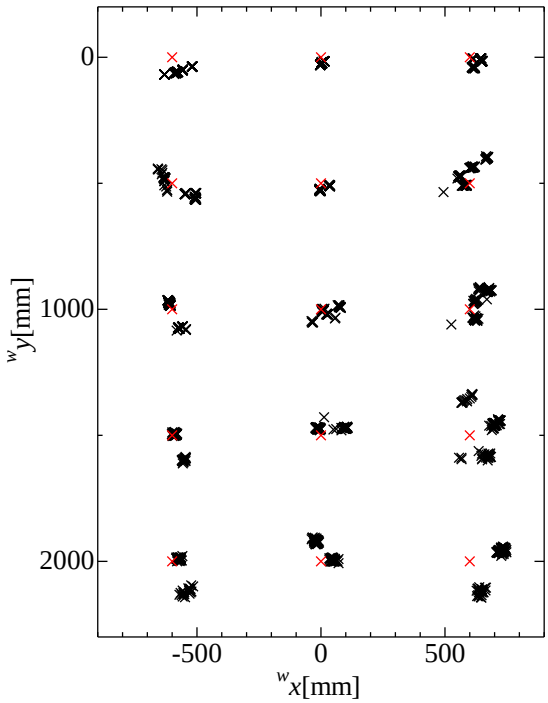

Figure 9. Experimental results with the pan–tilt camera (QCAM - 200RX, Logicool) set on ^{w}z=1350mm(×: actual position of the pan–tilt camera, ×: measurement result from the image obtained by the pan–tilt camera).

The error range as shown in Fig.9 means that it is possible to approach the multi-copter closer to a bridge girder to obtain clearer images for the inspection by utilizing the proposing method for applications such as operating based on the obtained position information and warning about an abnormal approach to a girder.

5. Conclusion

In this study, we proposed a self-position measurement method that employs the vertices of an image obtained by attaching a pan–tilt camera to a multi-copter and photographing an equilateral triangle landmark and estimates

how the method affects the results when extracting the vertices from the image. In addition, to suppress the measurement errors to a practical level, even when a multi-copter photographs a landmark from any direction, we tested a 3D tetrahedron as a landmark. Multi-copter photography makes a foothold unnecessary, and we can expect significant cost reduction. In addition, the results obtained in this study can be applied to the automatic navigation of multi-copters when GPS cannot be used, for example, in tunnels and buildings. Moreover, because multi-copters move at low speeds and take pictures at regular intervals, we can easily combine the images to study the conditions under the girder at a glance. The VR data created in this manner not only help to find abnormal locations, such as cracks in concrete, but also support the visual inspection of bridges, storage of inspection data, study of inspection technology, examination of inspection results, and evaluation of the aging process based on long-term data.

References:

[1] A. Miyamoto, H. Emoto, and H. Asano, Advanced performance evaluation system for existing concrete bridges, Computers and Concrete, Vol. 14(6), 2014, pp.727-743.

[2] G. Schutter, Advanced monitoring of cracked structures using video microscope and automated image analysis, NDT&E International, Vol. 35, 2002, 209-212.

[3] T. Yamaguchi and S. Hashimoto, Fast crack detection method for large-size concrete surface images using percolation-based image processing, Machine Vision and Applications, Vol. 21, 2009, pp.797-809.

[4] R. Adhikari, O. Moselhi, and A. Bagchi, Image-based retrieval of concrete crack properties for bridge inspection, Automation in Construction, Vol. 39(1), 2014, pp.180-194.

[5] J. Valença, D. Dias-da-Costa, E. Júlio, H. Araújo, and H. Costa, Automatic crack monitoring using photogrammetry and image processing, Measurement, Vol. 46, 2013, pp.433–441.

[6] J. Oh, G. Jang, S. Oh, J. Lee, B. Yi, Y. Moon, J. Lee, and Y. Choi, Bridge inspection robot system with machine vision, Automation in Construction, Vol. 18(7), 2009, pp.929-941.

[7] G. Li, S. He, Y. Ju, and K. Du, Long-distance precision inspection method for bridge cracks with image processing, Automation in Construction, Vol. 41, 2014, pp.83-95.

[8] N. Imajo, Y. Takada, and M. Kashinoki, Development and evaluation of compact robot imitating a hermit crab for inspecting the outer surface of pipes, Journal of Robotics, Vol.2015, Article ID 312780, 2015, 7 pages.

[9] N. Pham, H. La, Q. Ha, S. Dang, A. Vo, and Q. Dinh, Visual and 3D Mapping for Steel Bridge Inspection Using a Climbing Robot, 33rd International Symposium on Automation and Robotics in Construction and Mining, Auburn, Alabama, USA, 2016, pp.125-132.

[10] Y. Takada, S. Ito, and I. Imajo, Development of a Bridge Inspection Robot Capable of Traveling on Splicing Parts, Inventions, Vol. 2(3), 2017, DOI: 10.3390.

[11] R. Wang and Y. Kawamura, An Automated Sensing System for Steel Bridge Inspection Using GMR Sensor Array and Magnetic Wheels of Climbing Robot, Journal of Sensors, Vol.2016. Article ID 8121678, 2016, 15 pages.

[12] K. Sato, An Autonomous Flight Control of Multirotor Copter Using Video Image Information, Journal of the Society of Instrument and Control Engineers, Vol. 5(1), 2017, pp.32-35 (in Japanese).

[13] S. Ogido, A. Kyan, S. Takazato, R. Maezato, S. Tansuriyavong, and T. Anezaki, Proposed Integrated Drone Navigation and Autonomous Flight System for Aerial Inspection and Surveillance of Infrastructure Objects, IEEJ Transactions on Industry Applications, Vol. 136(10), 2016, pp.753-759 (in Japanese).

[14] W. J. Wolfe, D. Mathis, C.W. Sklair, and M. Magee, The perspective view of three points, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13(1), 1991, pp.66-73.

[15] R. Haralick, C. Lee, K. Ottenberg, and M. Nolle, Review and analysis of solutions of the three-point perspective pose estimation problem, International Journal of Computer Vision, Vol.13(3), 1994, pp.331-356.