AJG

Article

Adjustment of Alignment Errors of Disparity Images Around a Crane

K. Kazama *[1], J. Susaki[2] and Y. Ishii[3]

[1]Graduate School of Engineering, Kyoto University, Kyotodaigaku Katsura, Nishikyoku, Kyoto 615-8540, Japan. Email: <kazama.kojiro.34n@st.kyoto-u.ac.jp>

[2]Professor, Graduate School of Engineering, Kyoto University, Kyotodaigaku Katsura, Nishikyoku, Kyoto 615-8540, Japan. Email:<susaki.junichi.3r@kyoto-u.ac.jp>

[3]Assistant Professor, Graduate School of Engineering, Kyoto University, Kyotodaigaku Katsura, Nishikyoku, Kyoto 615-8540, Japan. Email:<ishii.yoshie.4k@kyoto-u.ac.jp>

*Corresponding author: K. Kazama, Email: <kazama.kojiro.34n@st.kyoto-u.ac.jp>

Received: November 21, 2024; Accepted: February 18, 2025; Published: March 25, 2025

ABSTRACT

Automated construction sites require rapid three-dimensional mapping of crane surroundings. Three-dimensional maps are generated using disparity images from a monocular camera mounted on a crane boom. Existing methods use homography matrices repeatedly to align disparity images, which accumulates errors and causes a misalignment in the origin position. In this study, a new alignment method was developed to address this challenge using the features of the crane orbit, such as the fixed base, boom length, and elevation angle. This approach uses rotation angles and swing radii estimated from homography matrices with adjustment and optimization through loop closure of the round-trip orbit of a crane. Parameter adjustments were based on fixed-point and angle comparisons between outward and return trips. The optimization was performed using the least-squares method. Experiments using round-trip images generated by the simulator yields reduced alignment errors. In future studies, we intend to verify whether the proposed method can be applied to real round-trip video images.

Keywords: Photogrammetry, Disparity image, Alignment error, Three-dimensional map

1. INTRODUCTION

Recently, chronic labor shortages at construction sites have become a problem. Thus, it is necessary to improve labor productivity and efficiency. Automation of crane operations has attracted attention as a solution to this challenge. For the automation of crane operations, a three-dimensional map of a construction site is required for the crane system to recognize the surroundings and determine safe lifting routes for loading (Lin et al., 2023). In general, three-dimensional maps are generated using lasers or moving images (Emmanuel, 1999). However, despite the decrease in the cost of light detection and ranging in recent years, the sensor component still favors cameras in terms of cost. Therefore, previous studies focused on generating the three-dimensional map of the surroundings of a crane using moving images acquired from a monocular camera. For example, Shigemori and Susaki

(2023) developed a method that used ORB-SLAM to generate three-dimensional maps. Shigemori et al. (2024) developed a method for the automatic classification of feature points used in three-dimensional mapping around slewing and derricking cranes. Kobayashi (2024) established a method to generate a real-scale three-dimensional orthographic map of the surroundings of a crane using disparity images generated from moving images. Three-dimensional maps can be plotted with a relatively high degree of accuracy using the method proposed by Kobayashi (2024).

In the existing method, a three-dimensional map is generated by integrating several disparity images obtained by rectifying the frame images and calculating the disparity for each pixel. For rectification, homography matrices between frames are used repeatedly to unify the coordinate system and align disparity images. However, existing alignment methods still have a problem. By using homography matrices repeatedly, alignment errors accumulate, causing the origin to be distant from the actual position.

Therefore, the objective of this study was to develop a new alignment method for disparity images to adjust the accumulated errors associated with the alignment of disparity images. The proposed approach uses a new alignment method that considers the features of a crane orbit, such as the fixed base, boom length, and apparent elevation angle of the crane. During the operation, the elevation angle changes steadily, but the desired modus operandum is to minimize such movements. Boom length is one of the main parameters that influence the maximum load; therefore, despite possible changes, it typically remains constant. Based on these features, we used parameters such as rotation angle and swing radius for alignment in this study. We estimated these parameters from the homography matrices by adjusting and optimizing them using the loop closure of the round-trip orbit of the crane.

2. EXISTING METHOD

2.1 Flow to generate disparity images

According to Kobayashi (2024), disparity images around a crane are generated using the following steps for the input moving images: (1) frame extraction and unit division, (2) determination of features and tie points, (3) image rectification, (4) generation of disparity images, (5) adjustment of disparity images, (6) projective transformation, and (7) generation of height images. In process (1), frames are extracted from the input moving images at equal intervals. The extracted frames are divided into several units to use more frames from the moving images and ensure that the rotation angles between frames are sufficiently large for the stable calculation of the homography matrices. In process (2), feature points, which are points in an image with local features, such as edges and intensity gradients, are extracted from the images, and the points extracted from the same object are determined as tie points based on the feature values. In process (3), image rectification is performed by applying homography matrices calculated using the tie points in pairs of images. In process (4), disparity images are generated by calculating the disparity for each pixel using semiglobal matching (Hirschmuller, 2008). The disparity is the distance in the image plane between a pair of tie points in two images. In this method, the disparity of the plane with the most feature points used for calculating the homography matrix for

rectification is zero. In process (5), adjustments such as offset calculation to integrate the reference plane and scaling of disparity values are conducted. In process (6), projective transformations of disparity images are performed based on the geometric relationships between disparities and locations on a plane to generate an orthographic map. In process (7), the disparity images are repeatedly aligned to an integrated coordinate system using a homography matrix, and the disparity images are translated into height images. Because the objective of this study was to adjust the alignment error, we mainly treated the processes related to the alignment of disparity images.

2.2 Existing Alignment Method

When the feature points of image i are expressed as (x_i, y_i), the homography matrix H_i is calculated as follows.

$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = H_i \begin{pmatrix} x_{i+1} \\ y_{i+1} \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_{i+1} \\ y_{i+1} \\ 1 \end{pmatrix} \quad (1)$$

Because this matrix has an indefinite scale, we set $h_{33} = 1$. Equation (1) can be transformed into Equation (2).

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} x_{i+1} & y_{i+1} & 1 & 0 & 0 & 0 & -x_i x_{i+1} & -y_i x_{i+1} \\ 0 & 0 & 0 & x_{i+1} & y_{i+1} & 1 & -x_i y_{i+1} & -y_i y_{i+1} \end{pmatrix}$$

$$(h_{11} \quad h_{12} \quad h_{13} \quad h_{21} \quad h_{22} \quad h_{23} \quad h_{31} \quad h_{32})^T \quad (2)$$

Assuming that three-dimensional points exist on the same plane, a homography matrix H_i can be obtained if at least four tie points exist. Disparity images are calculated by rectifying images using these homography matrices, which are also used to align disparity images, because homography matrices match points in one image to those in another. When disparity image i is DI_i, and DI_i aligned to DI_1 is denoted DI_i', DI_i' can be described as follows.

$$DI_i' = H_1 H_2 \cdots H_{i-1} DI_i \quad (3)$$

The coordinates of each disparity image are integrated by aligning all disparity images with the first disparity image.

3. PROPOSED METHOD

The orbit of the monocular camera attached to the tip of the crane boom has several features. For example, the center point is fixed when the crane base does not move, and the swing radius is fixed when the crane boom length and elevation angle are the same throughout the operation. In this paper, we propose a new alignment method that uses these features.

3.1 Proposed Alignment Method

Because disparity images are calculated from the moving images captured from the monocular camera mounted at the tip of the crane boom swinging at the same radius, disparity images can be aligned using the rotation radius and rotation angle between image i and image $i+1$, which are denoted r and θ_i in Figure 1, respectively.

Thus, when the image height and width are expressed as h and w, respectively, the alignment matrix T_i, which aligns x_i coordinates to $X-Y$ coordinates, is as Eq. (4).

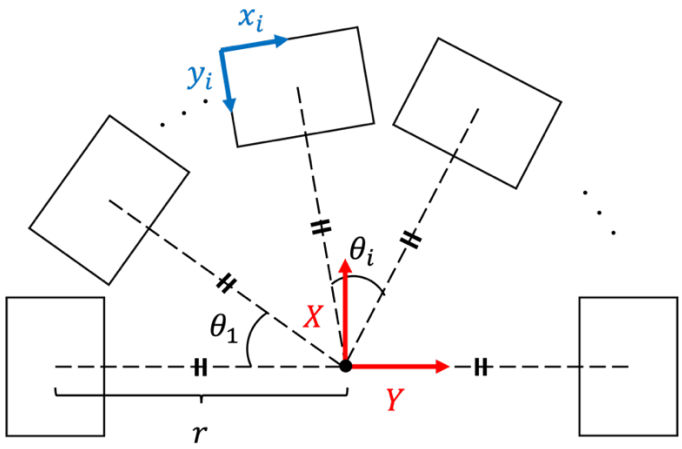

Figure 1: Variables used for optimization.

$$T_i = \begin{pmatrix} \cos\sum_{j=1}^{i-1}\theta_j & -\sin\sum_{j=1}^{i-1}\theta_j & 0 \\ \sin\sum_{j=1}^{i-1}\theta_j & \cos\sum_{j=1}^{i-1}\theta_j & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\frac{w}{2} \\ 0 & 1 & r-\frac{h}{2} \\ 0 & 0 & 1 \end{pmatrix} \tag{4}$$

When disparity image i is DI_i, and DI_i aligned to DI_1 is denoted DI_i', DI_i' can be described as follows.

$$DI_i' = T_i DI_i \tag{5}$$

Similarly, the units are aligned using the rotation angle estimated from the homography matrix calculated from the first image of each unit.

3.2 Calculating parameters from homography matrix

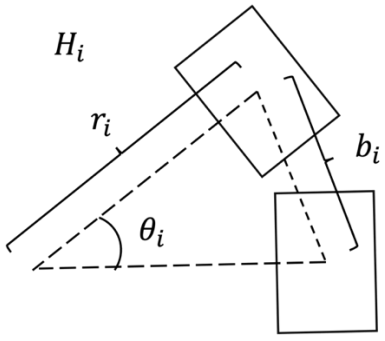

Figure 2: Variables from homography matrix.

From the homography matrix H_i, the distance between principal points b_i and rotation angle θ_i can be estimated. Each parameter is shown in Figure 2.

The distance between the principal points b_i is calculated by applying a homography matrix to the principal point. The rotation angle, θ_i, can be calculated as follows. At here, it is assumed that the influence of the scaling element is minimal because the camera height is almost constant and that the translational and rotational elements account for most of the homography matrices calculated using moving images captured from a swinging crane. Furthermore, the shear elements are also considered to better apply the effect of different viewpoints, so that homography matrix H_i can be expressed as follows. When the x- and y-translational components are t_x and, t_y respectively, the rotational component is θ_i, and the x- and y-shearing components are h_x and h_y, respectively.

$$H_i = \begin{pmatrix} 1 & h_x & 0 \\ h_y & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_i & -\sin\theta_i & t_x \\ \sin\theta_i & \cos\theta_i & t_y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta_i + h_x \sin\theta_i & -\sin\theta_i + h_x \cos\theta_i & t_x + h_x t_y \\ h_y \cos\theta_i + \sin\theta_i & -h_y \sin\theta_i + \cos\theta_i & h_y t_x + t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (6)$$

When the elements of homography matrix H_i calculated directly from the images are a, b, c, d, e, f, g, h, and 1, the equation is as follows.

$$H_i = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta_i + h_x \sin\theta_i & -\sin\theta_i + h_x \cos\theta_i & t_x + h_x t_y \\ h_y \cos\theta_i + \sin\theta_i & -h_y \sin\theta_i + \cos\theta_i & h_y t_x + t_y \\ 0 & 0 & 1 \end{pmatrix} \quad (7)$$

Thus, each parameter is formulated using a least-squares method.

$$\theta = \tan^{-1}\frac{ah_x - b + d - eh_y}{a + bh_x + dh_y + e}, h_x = \frac{a\sin\theta + b\cos\theta - t_x t_y + ct_y}{t_y^2 + 1}$$
$$h_y = \frac{d\cos\theta - e\sin\theta - t_x t_y + ft_y}{t_x^2 + 1}, t_x = \frac{-h_x t_y + c - h_y t_x + fh_y}{h_y^2 + 1} \quad (8)$$
$$t_y = \frac{-h_x t_x + ch_x - h_y t_y + f}{h_x^2 + 1}$$

By applying an iterative method (Kelley, 1995) to Equation (8), the rotation angle θ can be estimated. The swinging radius, r_i, is calculated as $r_i = b_i/\theta_i$ when θ_i is sufficiently small.

3.3 Optimization of parameters

Each rotation angle calculated from the corresponding homography matrix includes a small error, resulting in an accumulated error for the alignment based on Equation (4). Therefore, parameter optimization is necessary. When the observed and calculated rotation angles and swing radii calculated from the homography matrix H_i are denoted φ_i and r_i, respectively, residuals ε_i and δ_i will as expressed as follows. Here, θ_i and r are the true rotation angle and swing radius, respectively.

$$\varepsilon_i = \varphi_i - \theta_i, \delta_i = r_i - r \ (i = 1, 2, \cdots, n) \quad (9)$$

For optimization, we set a constraint condition as follows:

$$\sum_{i=1}^{n} \theta_i = \varphi_{1n} \quad (10)$$

where φ_{1n} is the rotation angle estimated from a homography matrix calculated directly from the first and last images. Therefore, the objective function is expressed as follows, using the least-squares method.

$$\min \sum_{i=1}^{n} \varepsilon_i^2 + \sum_{i=1}^{n} \delta_i^2$$
$$subject\ to: \sum_{i=1}^{n} \theta_i = \varphi_{1n} \quad (11)$$

The problem is formulated as follows using the Lagrange multiplier principle (Morton, 1950).

$$L = \frac{1}{2}\left(\sum_{i=1}^{n} \varepsilon_i^2 + \sum_{i=1}^{n} \delta_i^2\right) - \lambda\left(\varphi_{1n} - \sum_{i=1}^{n} \theta_i\right) \quad (12)$$

$$\begin{cases} \dfrac{\partial L}{\partial \theta_i} = -(\varphi_i - \theta_i) + \lambda = 0 \\ \dfrac{\partial L}{\partial r} = -\sum_{i=1}^{n}(r_i - r) = 0 \\ \dfrac{\partial L}{\partial \lambda} = -\left(\varphi_{1n} - \sum_{i=1}^{n}\theta_i\right) = 0 \end{cases} \quad (13)$$

Therefore, each variable is expressed as follows.

$$\lambda = \frac{\varphi_{1n}}{n} - \frac{\sum_{i=1}^{n}\varphi_i}{n}, r = \frac{\sum_{i=1}^{n} r_i}{n},$$

$$\theta_i = \varphi_i + \frac{\varphi_{1n}}{n} - \frac{\sum_{i=1}^{n}\varphi_i}{n} \ (i = 1, 2, \cdots, n) \quad (14)$$

This implies that the optimized radius is obtained by determining the mean, and the optimized rotation angles are obtained by equally adjusting the error closure.

3.4 Adjustment of parameters prior to optimization

Because each homography matrix is calculated from the tie points of a pair of images, it sometimes includes a significant error when matching fails. This error affects the accuracy of the optimization explained above; therefore, an adjustment is necessary. For this adjustment, the round-trip orbit of a crane is used. When fixed points are set according to the similarity of images on outward and return trips, the rotation angles between the fixed points in both trips are nearly identical. If the rotation angle between them differs significantly, the rotation angles are adjusted on the return trip.

When Im_t and Im_u are the images on the return trip corresponding to the images on outward trips Im_k and Im_j, respectively, as shown in Figure 3, the adjustment is as follows where $\Delta\Theta_{j,k}$ is the difference between the sum of rotation angles between fixed points.

$$\Delta\Theta_{j,k} = \sum_{i=j}^{k-1}\theta_i + \sum_{s=t}^{u-1}\theta_s \quad (15)$$

When the mean and standard deviation of the difference of the rotation angles between all fixed points $\Delta\Theta_{j,k}$ are described as $\mu_{\Delta\Theta}$ and $\sigma_{\Delta\Theta}$, the adjustment is performed as follows for every rotation angle between the fixed points in the return trip θ_s $(s = t, t+1, \cdots, u-1)$. At here θ_s' represent the rotation angles after adjustment.

$$\theta_s' = \begin{cases} \theta_s, & (\left|\Delta\Theta_{j,k} - \mu_{\Delta\Theta}\right| < 2\sigma_{\Delta\Theta}) \\ \theta_s - \dfrac{\Delta\Theta_{j,k}}{u-t}, & (\left|\Delta\Theta_{j,k} - \mu_{\Delta\Theta}\right| \geq 2\sigma_{\Delta\Theta}) \end{cases} \quad (16)$$

When a difference between the sum of the rotation angles between fixed points is an outlier according to the means and standard deviation, an adjustment is conducted by adjusting the error closure between fixed points to rotation angles in the return trip, assuming that the rotation angle in the outward trip is correct.

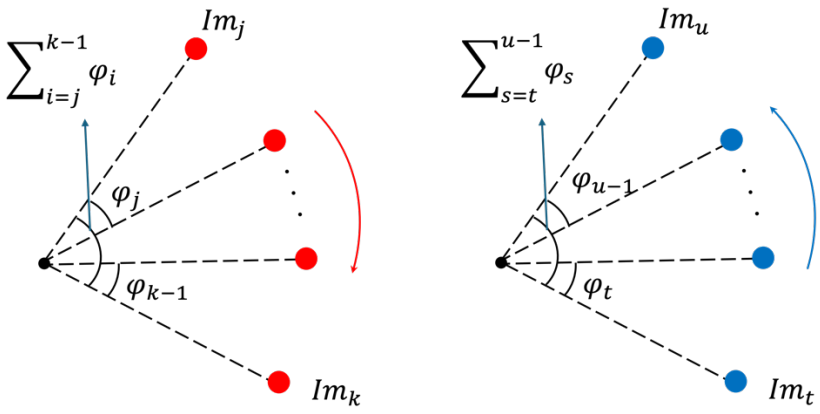

Figure 3: Adjustment of variables using fixed points. The left part represents the outward trip, and the right part represents the return trip.

4. DATA

Moving images were acquired by a simulator using the Gazebo software (Open Robotics, 2024). In the simulator, a camera model was mounted on a crane model that imitated an actual crane, and a building model with a height of 3.6–4.0 m was built in the environment. The crane rotated by approximately 140° at a speed of 7 °/s. The simulated camera height and radii were approximately 15.2 m and 7.4 m. These values are derived from the camera position obtained in the simulator. The 140° crane rotation parameter is selected in order to emphasize and focus on the evaluation of the building height and location because the moving images taken in the other 220° only shows the ground part. Moving images were captured at a resolution of 1280×720 pixels at 30 fps, and they consisted of 609 frames of approximately 20 s. Figure 4 shows an example of the images captured by the simulator.

As these moving images only made a half-turn, we created a round-trip orbit by adding the images obtained during the outward trip to the return trip. Figure 5 illustrates the concept of folding back the images obtained during the outward trip to the return trip. Because the images were identical in both trips, this corresponded to $\varphi_{1n} = 0$ in Equation (14) and $k = j + 1$, $u = t + 1$ in Equations (15) and (16), respectively.

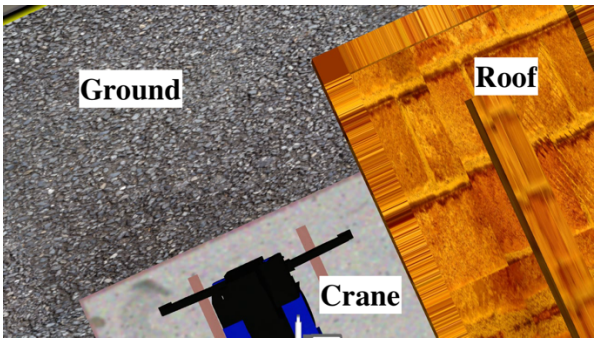

Figure 4: Moving Images taken from a gazebo simulator.

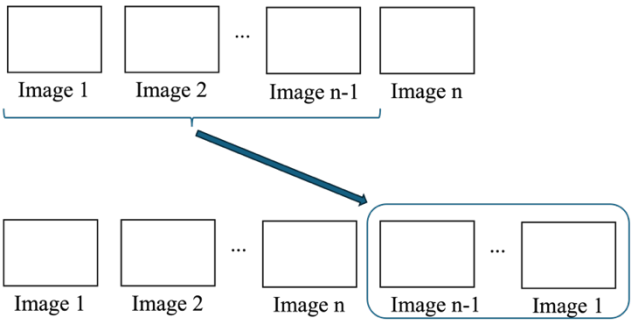

Figure 5: Folding back of moving images.

5. RESULTS

In the experiment, 60 frames were extracted from 609 frames of moving images and divided into three units of 20 frames per unit. After the division, 19 frames were added to the unit by folding them back. Consequently, 38 disparity

images were generated for each unit. In this section, we present comparisons with the alignment results obtained using each method for images with the same disparity.

5.1 Alignment using homography matrices

Figure 6 shows the final height images generated by integrating all the units.

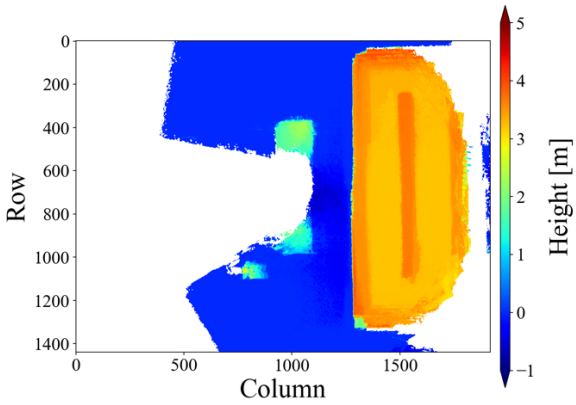

Figure 6: Final height images generated by integrating all units aligned by homography.

The transformation matrices used for the alignment of the last disparity image to the first disparity image for unit 0, 1, 2 were

$$\begin{pmatrix} 1.01 & -2.52 \times 10^{-3} & 1.56 \times 10 \\ 1.04 \times 10^{-2} & 9.81 \times 10^{-1} & 2.05 \times 10 \\ 1.56 \times 10^{-5} & 2.23 \times 10^{-6} & 9.86 \times 10^{-1} \end{pmatrix},$$

$$\begin{pmatrix} 9.94 \times 10^{-1} & -6.04 \times 10^{-4} & -1.02 \times 10 \\ -1.21 \times 10^{-2} & 1.02 & 3.01 \times 10 \\ -1.08 \times 10^{-5} & -1.04 \times 10^{-6} & 1.01 \end{pmatrix},$$

$$\begin{pmatrix} 9.85 \times 10^{-1} & -4.16 \times 10^{-3} & -1.14 \\ -2.09 \times 10^{-2} & 1.01 & 2.22 \times 10 \\ -1.83 \times 10^{-5} & -2.01 \times 10^{-6} & 1.03 \end{pmatrix}.$$

5.2 Alignment using rotation angles and swing radius with adjustment and optimization

Figure 7 shows the final height images generated by integrating all the units. The sum of the rotation angles calculated from the homography matrices for units 0, 1, and 2 were 1.07×10^{-1}, 9.46×10^{-3}, and 2.69×10^{-2} rad, respectively. In the experiment, all frames used during the outward and return trips were the same; thus, all frames could be adopted as fixed points. This corresponds to $k = j + 1, u = t + 1$ in Equations (15) and (16), respectively. Thus, 20 fixed points were in each unit. The rotation differences between the fixed points for units 0, 1, and 2 are listed in Tables 1, 2, and 3, respectively, where the adjusted angles are in red.

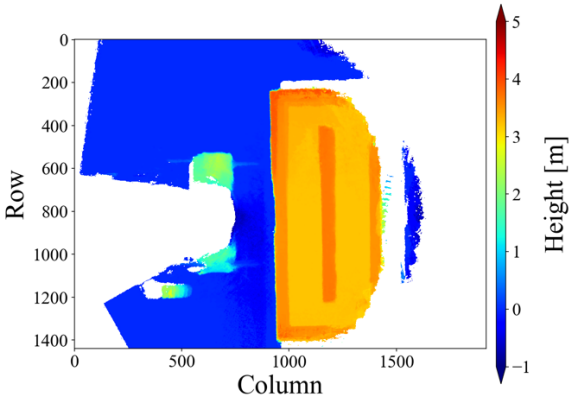

Figure 7: Final height images generated by integrating all units aligned by rotation angles and swing radius.

Table 1: Rotation differences between fixed points for unit 0.

i	1	2	3	4	5
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-1.95	2.44	5.25	0.122	-8.67
i	6	7	8	9	10
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-0.122	1.59	18.7	24.3	9.52
i	11	12	13	14	15
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-8.55	-0.855	-0.855	-8.55	-4.27
i	16	17	18	19	
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	29.5	2.20	28.3	18.4	

Table 2: Rotation differences between fixed points for unit 1.

i	1	2	3	4	5
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-7.20	-3.05	-1.34	4.03	-4.76
i	6	7	8	9	10
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	5.49	11.7	33.1	6.23	1.22
i	11	12	13	14	15
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	0.488	-10.5	-19.4	-31.1	0.610
i	16	17	18	19	
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-1.71	0.610	27.8	-2.81	

Table 3: Rotation differences between fixed points for unit 2.

i	1	2	3	4	5
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-4.64	6.90	134	0.122	-0.244
i	6	7	8	9	10
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	26.0	38.3	-0855	-53.8	-7.32
i	11	12	13	14	15
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	-3.54	27.3	7.32	2.56	7.81
i	16	17	18	19	
$\Delta\Theta_{i,i+1}[\text{rad}\times 10^{-3}]$	12.6	24.9	-3.17	0.00	

Table 4: Transitions of differences of rotation sum from the true rotation sum.

Unit No.	Unit 0	Unit 1	Unit 2
Initial	1.07×10^{-1}	9.46×10^{-3}	2.69×10^{-2}
After adjustment	1.07×10^{-1}	7.45×10^{-3}	8.07×10^{-2}
After optimization	-6.10×10^{-6}	-1.65×10^{-3}	1.59×10^{-3}

Thus, the means of the rotation differences between the fixed points $\Delta\Theta_{j,k}$ for units 0, 1, and 2 were 5.61×10^{-3}, 4.98×10^{-4}, and 1.41×10^{-3} rad, respectively. The standard deviations of the rotation differences between the fixed points $\Delta\Theta_{j,k}$ for units 0, 1, and 2 were 1.20×10^{-3}, 1.38×10^{-2}, and 1.89×10^{-2} rad, respectively. Thus, according to Equation (16) and Table 1, 2 and 3, the no rotation angle in unit 0, two rotation angles in unit 1, and one rotation angle in unit 2 were adjusted. After the adjustment, the sums of the rotation angles for units 0, 1, and 2 were 1.07×10^{-1}, 7.45×10^{-3}, and 8.07×10^{-2} rad, respectively. By dividing these sums with the number of the angles, which corresponds to $\sum_{i=1}^{n} \varphi_i / n$ in Equation (14), each rotation angle in each unit was adjusted with -2.81×10^{-3}, -1.96×10^{-4}, and -2.12×10^{-3} rad because the true rotation sum φ_{1n} is 0 rad in Equation (14). So that, the sum of rotation angles for units 0, 1, and 2 became -6.10×10^{-6}, -1.65×10^{-3}, and 1.59×10^{-3} rad, respectively, after the whole adjusting and optimization process. Table 4 shows these transitions of differences of rotation sum compared with the true rotation sum φ_{1n} in each unit.

By determining the mean of all swing radii calculated from the homography matrices, the integrated swing radius became $r = 4.64 \times 10^2$ pixels. Figure 8 shows an example of the differences between the height images of unit 0 with and without parameter optimization generated using images from the return trip.

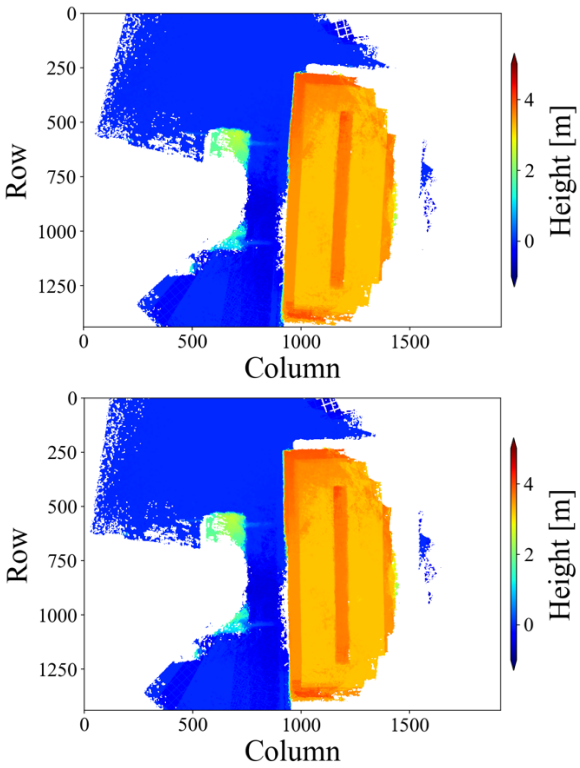

Figure 8: Height images generated by integrating images during the return trip in unit 0. (Top: w/o optimization. Bottom: with optimization.)

Height images based on only return trips are shown because the effect of accumulation errors in the alignments is more significant in the later part of the moving images. Furthermore, to optimize the rotation angles, we used the same

radius as in the height image without optimization.

Figure 9 shows the delta height map generated from height images showed in Figure 8. It is generated by subtracting the height image of return trip in unit 0 without optimization from that with optimization.

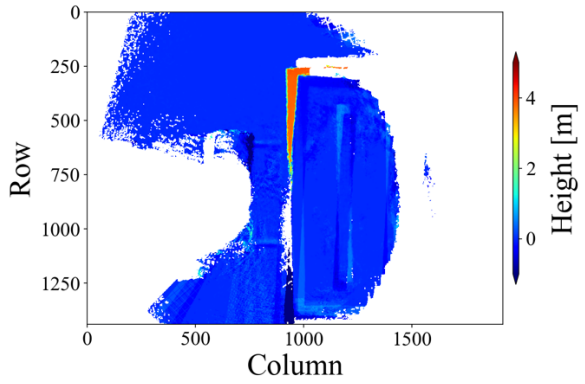

Figure 9: Delta height map generated by subtracting the height image of return trip in unit 0 without optimization from that with optimization.

6. DISCUSSION

In this paper, we propose a new alignment method using parameters, such as the rotation angles between images and swing radius. The parameters were optimized using the least-squares method from homography matrices, and the rotation angles were adjusted based on the differences between the outward and return trips of the crane orbit.

6.1 Alignment using rotation angles and swing radii

A comparison of the alignment using rotation angles and swing radius with that using homography matrices, as shown in Figures 6 and 7, indicates that similar results can be obtained by applying the alignment method without using homography matrices directly, but using parameters calculated from homography matrices. We believe that this is possible because of the features of the crane orbit, which is a fixed-rotation center point of the crane boom without crane base movement and with a fixed crane boom length. However, the error at the center point of the last disparity image is different when compared in detail. In this experiment, all frames used during the return trip were identical to those used during the outward trip. Thus, the true transformation matrix for the alignment from the last disparity image to the first is an identity matrix. However, when focusing on the translation elements of the alignment using homography matrices, the final disparity images in units 0, 1, and 2 had translations of approximately 2.58×10, 3.18×10, and 2.22×10 pixels, respectively, compared to the first disparity images. Because the sum of the rotation angles and swing radius of units 0, 1, and 2 were -6.10×10^{-6}, -1.65×10^{-3}, and 1.59×10^{-3} rad, respectively, and 4.64×10^{2} pixels. The final disparity images in units 0, 1, and 2 had translations of approximately 2.83×10^{-3}, 7.66×10^{-1}, and 7.38×10^{-1} pixels, respectively, based on the alignment using rotation angles and swing radius. In the moving images, the camera height h was 15.243 m, the focal length of the camera was 931.205 pixels, and 1 pixel on the plane of the final height image was equal to 1.61×10^{-2} m as the real length. Thus, the accumulation error of the alignment in units 0, 1, and 2 improved from 4.15×10, 5.12×10, and 3.57×10 cm to 4.56×10^{-3}, 1.23, and 1.19 cm, respectively, via alignment using rotation angles and swing radius with adjustment and optimization. However, the true simulated radii on the image plane will be about 453.6 pixels because the simulated camera radius

is about 7.4 m, and the ratio of the focal length and the camera height is about 61.3 pixels/m. The difference is about 9.4 pixels compared with 464 pixels, and this corresponds to 0.153 m. This error might indicate that just taking the means of all radii might not be the best way to estimate the true radius, or the radius from homography matrix tends to be overestimated with the current method. Considering this, the alignment using these parameters is effective because of the constraints of the crane orbit, but the optimization and estimation method of parameters needs to be improved.

6.2 Optimization of parameters

Because all frames were the same during the outward and return trips, the sum of the rotation angles of all images should be 0 rad; hence, we compared the effect of the optimization using the sum of the rotation angles and the location of the center point of the last disparity image. The sums of the rotation angles and swing radius before the optimization of units 0, 1, and 2 were 1.07×10^{-1}, 7.45×10^{-2}, and 8.07×10^{-2} rad, respectively, and 464 pixels, and the final disparity images in units 0, 1, and 2 had translations of approximately 4.96×10, 3.46, and 3.74×10 pixels, respectively. Those after the optimization of units 0, 1, and 2 were -6.10×10^{-6}, -1.65×10^{-3}, and 1.59×10^{-3} rad, respectively, and 464 pixels, and the final disparity images in units 0, 1, and 2 had translations of approximately 2.83×10^{-3}, 7.66×10^{-1}, and 7.38×10^{-1} pixels, respectively. Figure 8 shows the height images with and without parameter optimization using images captured during the return trip of unit 0. We only show the image of unit 0 because the change is larger than that in the other units. The rotation error in the left image was adjusted to the right image via optimization. Thus, the accumulated errors in the sum of the rotation angles and the estimated location of the last disparity image are improved.

6.3 Adjustment of parameters

Immediately after the adjustment using fixed points, the sum of the rotation angles of units 0, 1, and 2 changed from 1.07×10^{-1}, 9.46×10^{-3}, and 2.69×10^{-2} rad to 1.07×10^{-1}, 7.45×10^{-2}, and 8.07×10^{-2} rad, respectively. In units 0 and 1, the difference of the rotation sums from the true sum zero does not changes or becomes smaller, but in unit 2 it becomes larger, from 2.69×10^{-2} to 8.07×10^{-2}.. Because no significant difference existed in the means and standard deviations of the rotation differences between the fixed points in the units, the adjustment limited to the return trip might be the reason. In this method, the rotation angle of the return trip is adjusted to that of the outward trip; however, it is impossible to determine which rotation angle of the trip includes errors. Thus, it is possible that the rotation angle of the outward trip is incorrect. For this adjustment, it may be necessary to consider which rotation angle should be adjusted by comparing the rotation angles between other fixed points.

7. CONCLUSION

Automatic crane operation has attracted attention in improving labor productivity at construction sites. A three-dimensional map of the crane surroundings is necessary for the crane system to recognize the surroundings and determine safe lifting routes of the loads. For the generation of three-dimensional maps, a method using disparity images generated from moving

images acquired from the tip of the crane boom using a monocular has been developed; however, this method yields an accumulation error in the alignment of each disparity image because it repeatedly uses homography matrices calculated from each pair of images for the alignment. To address this problem, we propose a new alignment method using rotation angles and swing radii with parameter adjustment and optimization processes. For the adjustment, we applied the differences in rotation angles between fixed points in the outward and return trips and used the least-squares method with residuals as rotation angles and swing radius calculated from homography matrices. The error from the true location of the last disparity image determined using the proposed alignment method is improved compared to the previous alignment error. Therefore, despite the limited use of the same image for outward and return trips, we developed a method to adjust the accumulation error in the alignment of disparity images. In the future, we aim to validate our current implementation of in situ data of a real crane round-trip motion and extend the method to account for changes in the elevation angle of the crane.

REFERENCES

Emmanuel P. B., (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry & Remote Sensing,* 54, 83-94.

Hirschmuller, H., (2008). Stereo processing by semiglobal matching and mutual information, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), 328-341.

Kazama, K., (2024). Developing integration method of disparity images around a crane using moving images. Japan: Graduation thesis, Kyoto University.

Kelley, C. T., (1995). Iterative methods for linear and nonlinear equations, North Carolina State University.

Kobayashi, T. (2024). Real-scale 3D ortho-mapping of the crane's surrounding from moving images considering the construction site. Japan: Master dissertation, Kyoto University.

Lin, X., Han, Y. & Guo, H. (2023). Lift path planning for tower cranes based on environmental points clouds. *Automation in Construction*, Volume 155, 105046. https://doi.org/10.1016/j.autcon.2023.105046

Modenov, P. S. & Parkhomenko, A. S. (2014). Euclidian and Affine Transformations: Geometric Transformations.

Morton, S. (1950). Lagrange multipliers revisited. *Cowles Commission Discussion Paper: Mathematics*, 403, November 7, 1950.

Open Robotics: GAZEBO, Retrieved on August 20, 2024, from https://gazebosim.org/home.

Shigemori, H., & Susaki, J., (2021). Realtime generation of 3D maps around cranes using ORB-SLAM2. *Committee of Infrastructure Planning and Management, 63rd research presentation and collection of papers* (in Japanese).

Shigemori, H., Susaki, J., Yoneda, M. & Ososinski, M., (2024). Development of an Automatic Feature Point Classification Method for Three-Dimensional Mapping Around Slewing and Derricking Cranes.

Photogrammetric Engineering & Remote Sensing, Vol. 90, No. 9, September 2024, pp. 537-551.