Adjustment of Alignment Error of Disparity Images Around a Crane

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

[1]Student, Graduate School of Engineering, Kyoto University, Japan

[2]Professor, Graduate School of Engineering, Kyoto University, Japan

[3]Assistant Professor, Graduate School of Engineering, Kyoto University, Japan

*kazama.koijro.34n@st.kyoto-u.ac.jp (*Corresponding author's email only)

Abstract *Automation at construction sites requires rapid generation of a three-dimensional map of the crane's surroundings. One approach for mapping is to use disparity images generated from moving images acquired by a monocular camera attached to the tip of a crane boom. In existing methods, disparity images are obtained by rectifying frame images and calculating disparity for each pixel. For rectification, homography matrices between frames are utilized, which are used repeatedly to unify the coordinate system of disparity images and align them. However, as a result of using homography matrices repeatedly, the accumulation of alignment errors appears, causing the origin position is away from the actual one. Our approach is to propose a new alignment method considering the features of crane orbit, like a fixed base, boom length and apparent elevation angle of the crane. During operations, the elevation angle changes readily, but the desired modus operandum is to minimize such movements. In regard to the boom length, it is one of the main factors affecting maximum load, therefore despite possible changes, it typically stays constant. In this paper, we utilized parameters such as rotation angles and swing radius estimated from homography matrices for the alignment, adjusting and optimizing those parameters using loop closure of a round-trip orbit of crane. Adjustment of parameters are performed by setting fixed points in the orbit and comparing the angles between outward trip and return trip. Optimization of parameters are conducted by using least squares method. For an experiment, we applied the method to moving images taken by a simulator, which have a round-trip orbit with completely same images in outward trip and return trip by reversely using images in outward trip as return trip. As a result, the alignment error was reduced by using the proposed method. Despite the limited situation of using same images for outward and return trips, it can be said that we were able to propose a method to compensate for an alignment error of disparity images. In the future, we would like to verify whether this method can be applied to real round-trip video images.*

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

Introduction

Chronic labor shortages at construction sites have recently become a problem. It is necessary to improve labor productivity and efficiency. To deal with this problem, automation of crane operations has attracted attention as a solution. For the automation of crane operations, three-dimensional map at a construction site is necessary for the crane system to recognize the surroundings and to determine safe lifting routes for loadings (Lin et al., 2023). In general, three-dimensional maps are generated using laser or moving images (Emmanuel, 1999). However, despite the decrease of the cost of the lidar in the recent years, the sensor component still favors cameras in terms of the cost. Therefore,

previous studies have been working on generating a three-dimensional map of a crane's surroundings using moving images acquired from a monocular camera. For example, Shigemori and Susaki (2023) developed a method using ORB-SLAM for generating a three-dimensional map. Furthermore, Shigemori et al. (2024) developed a method for automatic classification of feature points used in three-dimensional mapping around slewing and derricking cranes. Kobayashi (2024) exploited a method to generate a real-scale three-dimensional orthographic map of the crane's surrounding using disparity images generated from moving images. In particular, it is possible to generate a three-dimensional map in relatively high degree of accuracy using the method proposed by Kobayashi (2024).

In the existing method, a three-dimensional map is generated through integration of several disparity images, which are obtained by rectifying the frame images and calculating the disparity for each pixel. For rectification, homography matrices between frames are utilized, which are used repeatedly to unify the coordinate system and to align disparity images. However, there is still a problem with the existing alignment method. For now, as a result of using homography matrices repeatedly, the accumulation of alignment errors appears, causing the origin position is to be distant from the actual one.

Therefore, the objective of this paper is to develop a new alignment method for disparity images to adjust the accumulation errors associated with the alignment of disparity images. Our approach is to propose a new alignment method considering the features of crane orbit, like a fixed base, boom length and apparent elevation angle of the crane. During operations, the elevation angle changes readily, but the desired modus operandum is to minimize such movements. In regard to the boom length, it is one of the main factors affecting maximum load, therefore despite possible changes, it typically stays constant. According to these features, we utilized parameters such as rotation angles and swing radius for the alignment in this paper. We estimate these parameters from homography matrices, adjusting and optimizing them using loop closure of a round-trip orbit of crane.

Current Method

a. Flow to generate disparity images:

According to Kobayashi's method (2024), disparity images around a crane are generated using the following steps for input moving images: (1) Frame extraction and unit division,

(2) Determination of feature and tie points, (3) Image rectification, (4) Generation of disparity images, (5) Adjustment of the 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. In order to utilize more frames from the moving images and ensure that the rotation angles between frames are sufficiently large for stable calculation of the homography matrices, the extracted frames are divided into several units. In process (2), feature points, which are points in an image with local features such as edges and intensity gradients, are extracted from images, and the points extracted from the same object are determined as tie points according to feature values. In process (3), image rectifications are conducted using homography matrices calculated using tie points in pairs of images. In process (4), disparity images are generated by calculating the disparity for each pixel using semi-global matching (Hirschmuller, 2008). A 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 for calculating the homography matrix for rectification is obtained as zero. In process (5), adjustments such as the offset calculation to integrate the reference plane and the scaling of disparity values are conducted. In process (6), to generate orthographic map, projective transformations of disparity images are conducted according to geometric relationships with disparities and location on a plane. In process (7), disparity images are aligned to an integrated coordinate using homography matrix repeatedly, and disparity images are translated into height images. Since the objective of this paper is to adjust the alignment error, we will treat mainly about processes related to the alignment of disparity images from here.

b. Current Alignment Method:

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

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

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

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} x_{i+1} & y_{i+1} & 1 & 0 & 0 & 0 & -x_i x_{i+1} & -y_i x_{i+1} \\ 0 & 0 & 0 & x_{i+1} & y_{i+1} & 1 & -x_i y_{i+1} & -y_i y_{i+1} \end{pmatrix}$$
$$(h_{11} \quad h_{12} \quad h_{13} \quad h_{21} \quad h_{22} \quad h_{23} \quad h_{31} \quad h_{32})^T \#(2)$$

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

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

By aligning all disparity images to the first disparity image, coordinates of each disparity image will be integrated.

Proposed Method

The orbit of a monocular camera attached to the tip of crane boom has some features. For examples, 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 same through the operation. We propose a new alignment method using these features.

a. Proposed Alignment Method:

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

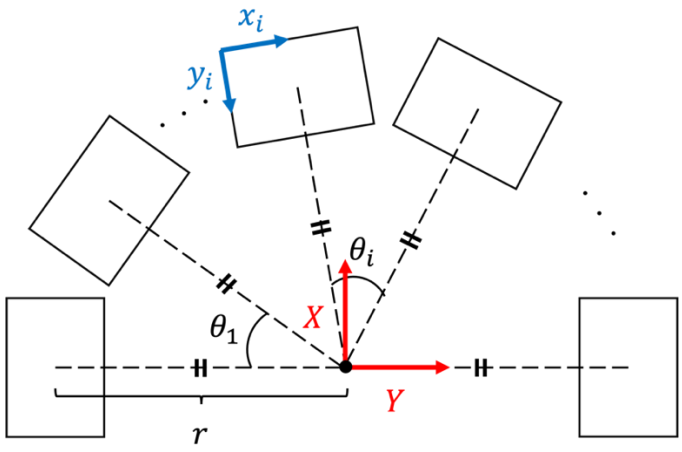

Figure 1: Variables used for an optimization.

So that, when image height and width are expressed as h and w, the alignment matrix T_i, which align x_i-y_i coordinate to X-Y coordinate, will be as follows.

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

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

$$DI_i' = T_i DI_i \#(5)$$

Similarly, the alignments between units are conducted by using rotation angle estimated from the homography matrix calculated from the first image of each unit.

b. Calculating parameters from homography matrix:

From the homography matrix $\boldsymbol{H_i}$, the distance between principal points $\boldsymbol{b_i}$ and the rotation angle $\boldsymbol{\theta_i}$ can be estimated. Each parameter is defined as Figure 2.

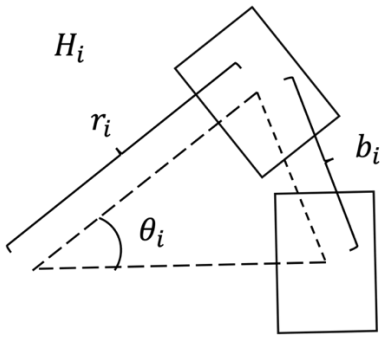

Figure 2: Variables from Homography Matrix.

The distance between the principal points $\boldsymbol{b_i}$ is calculated by applying a homography matrix to the principal point. The rotation angle $\boldsymbol{\theta_i}$ can be calculated as follows. By assuming that the influence of the scaling element is small and that the translational and rotational elements account for most of the homography matrices calculated using moving images taken from a swinging crane, a homography matrix $\boldsymbol{H_i}$ can be expressed as follows, when x and y translational components are $\boldsymbol{t_x}$ and $\boldsymbol{t_y}$, a rotational component is $\boldsymbol{\theta}$ and x and y shearing components are $\boldsymbol{h_x}$ and $\boldsymbol{h_y}$.

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

When the elements of homography matrix $\boldsymbol{H}$ calculated directly from the images are

$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h}$, 1, equation will be as follows.

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c} \\ \boldsymbol{d} & \boldsymbol{e} & \boldsymbol{f} \\ \boldsymbol{g} & \boldsymbol{h} & \mathbf{1} \end{pmatrix} = \begin{pmatrix} \cos\boldsymbol{\theta} + \boldsymbol{h_x}\sin\boldsymbol{\theta} & -\sin\boldsymbol{\theta} + \boldsymbol{h_x}\cos\boldsymbol{\theta} & \boldsymbol{t_x} + \boldsymbol{h_x t_y} \\ \boldsymbol{h_y}\cos\boldsymbol{\theta} + \sin\boldsymbol{\theta} & -\boldsymbol{h_y}\sin\boldsymbol{\theta} + \cos\boldsymbol{\theta} & \boldsymbol{h_y t_x} + \boldsymbol{t_y} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} \#(\mathbf{7})$$

Thus, each parameter will be as follows by formulating as a least squares method.

$$\boldsymbol{\theta} = \tan^{-1}\frac{\boldsymbol{a h_x} - \boldsymbol{b} + \boldsymbol{d} - \boldsymbol{e h_y}}{\boldsymbol{a} + \boldsymbol{b h_x} + \boldsymbol{d h_y} + \boldsymbol{e}}, \boldsymbol{h_x} = \frac{\boldsymbol{a}\sin\boldsymbol{\theta} + \boldsymbol{b}\cos\boldsymbol{\theta} - \boldsymbol{t_x t_y} + \boldsymbol{c t_y}}{\boldsymbol{t_y^2} + \mathbf{1}}$$
$$\boldsymbol{h_y} = \frac{\boldsymbol{d}\cos\boldsymbol{\theta} - \boldsymbol{e}\sin\boldsymbol{\theta} - \boldsymbol{t_x t_y} + \boldsymbol{f t_y}}{\boldsymbol{t_x^2} + \mathbf{1}}, \boldsymbol{t_x} = \frac{-\boldsymbol{h_x t_y} + \boldsymbol{c} - \boldsymbol{h_y t_x} + \boldsymbol{f h_y}}{\boldsymbol{h_y^2} + \mathbf{1}} \#(\mathbf{8})$$
$$\boldsymbol{t_y} = \frac{-\boldsymbol{h_x t_x} + \boldsymbol{c h_x} - \boldsymbol{h_y t_y} + \boldsymbol{f}}{\boldsymbol{h_x^2} + \mathbf{1}}$$

By using an iterative method (Kelley, 1995) to Equation (8), the rotation angle $\boldsymbol{\theta}$ can be estimated. The swinging radius $\boldsymbol{r_i}$ is calculated as $\boldsymbol{r_i} = \boldsymbol{b_i}/\boldsymbol{\theta_i}$ when $\boldsymbol{\theta_i}$ is small enough.

c. Optimization of parameters:

Each rotation angle calculated from the corresponding homography matrix includes a small error, resulting in an accumulation error for the alignment according to Equation (4). Therefore, an optimization of parameters is necessary. When the observed and calculated rotation angle and swing radius calculated from the homography matrix $\boldsymbol{H_i}$ are described as $\boldsymbol{\varphi_i}$ and $\boldsymbol{r_i}$, residuals $\boldsymbol{\varepsilon_i}$ and $\boldsymbol{\delta_i}$ will as stated in the following. At here, $\boldsymbol{\theta_i}$ and $\boldsymbol{r}$ are the true rotation angle and the true swing radius.

$$\boldsymbol{\varepsilon_i} = \boldsymbol{\varphi_i} - \boldsymbol{\theta_i}, \boldsymbol{\delta_i} = \boldsymbol{r_i} - \boldsymbol{r}\ (\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \cdots, \boldsymbol{n}) \#(\mathbf{9})$$

For optimization, we set a constraint condition as follows.

$$\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\theta_i} = \boldsymbol{\varphi_{1n}} \#(\mathbf{10})$$

At here, $\boldsymbol{\varphi_{1n}}$ is the rotation angle estimated from a homography matrix calculated directly from the first image and the last image. Therefore, the objective function is stated as follows using the least squares method.

$$\begin{aligned} &\mathbf{min} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\varepsilon_i^2} + \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\delta_i^2} \\ &\boldsymbol{subject\ to}: \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{\theta_i} = \boldsymbol{\varphi_{1n}} \end{aligned} \quad \#(\mathbf{11})$$

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

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

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

Thus, each variable is expressed as follows.

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

This means that optimized radius is obtained by taking the mean and optimized rotation angles are obtained by equally adjusting the error closure.

d. Adjustment of parameters before optimization:

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

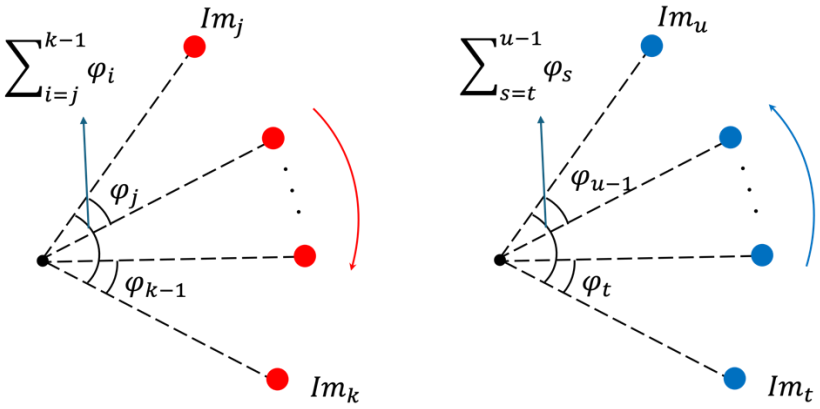

Figure 3: Variables adjustment using fixed points. Left figure is outward trip and right figure is return trip.

When Im_t and Im_u are the images on the return trip corresponding to the images on the outward trip Im_k and Im_j as shown in the Figure 3, the adjustment will be as follows.

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

At here, $\Delta\Theta_{j,k}$ is defined as a difference between the sum of rotation angles between fixed points. When the mean and the standard deviation of $\Delta\Theta_{j,k}$ between all fixed points are described as $\mu_{\Delta\Theta}$ and $\sigma_{\Delta\Theta}$, the adjustment will be as follows.

$$\theta_{ps} = \begin{cases} \theta_s & if \left|\Delta\Theta_{j,k} - \mu_{\Delta\Theta}\right| < 2\sigma_{\Delta\Theta} \\ \theta_s - \dfrac{\Delta\Theta_{j,k}}{u-t} & else \end{cases} \quad (s = t, t+1, \cdots, u-1) \#(16)$$

At here, when adjustment is conducted, the error closure is adjusted equally to rotation angles in return trip.

Data

Moving images were acquired by a simulator using 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 installed in an environment. The crane rotated about 140 deg at a speed of 7 deg/s. The camera height was approximately 15 m. Moving images were obtained at a resolution of 1280×720 pixels at 30 fps and consist of 609 frames of approximately 20 seconds. Figure 4 shows an example image of the simulator.

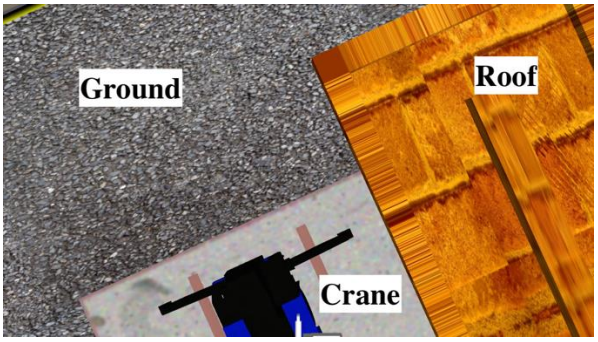

Figure 4: Moving Images taken from a gazebo simulator.

Since this moving images only make a half turn, we created a round-trip orbit by adding the images obtained in the outward trip as the return trip. Figure 5 shows the concept of folding back the images obtained in the outward trip to return trip. Since used images are identical in both trip, this corresponds to $\varphi_{1n} = 0$ in Equation (14) and $k = j + 1,\ u = t + 1$ in Equation (15) and (16).

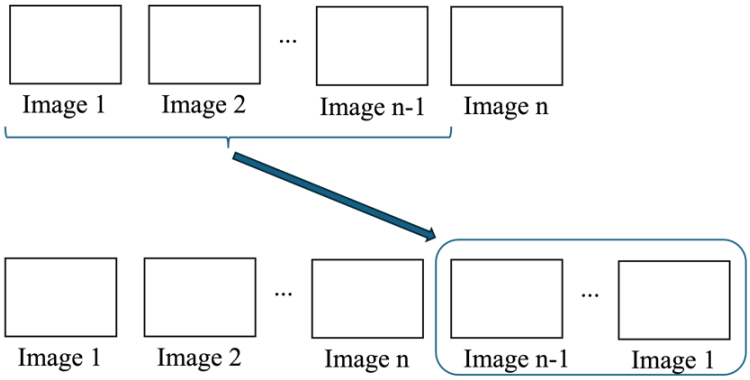

Figure 5: Folding Back of Moving Images.

Results

In this experiment, 60 frames were extracted from 609 frames of the moving images and divided into 3 units of 20 frames in each unit. After division, 19 frames were added into the unit by folding back the frames. As a result, 38 disparity images were generated for each unit. In this section, we compare the alignment results obtained by each method using the same disparity images.

a. Alignment using homography matrices:

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

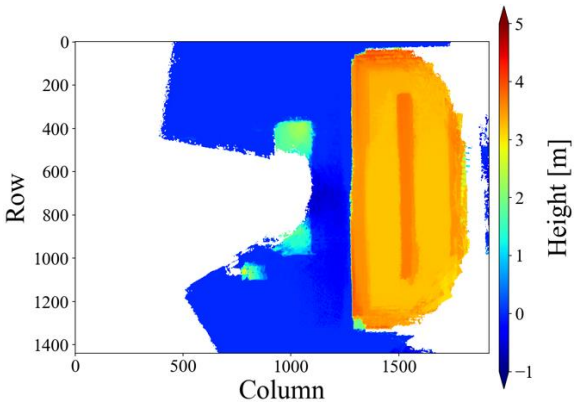

Figure 6: Final height images generated by integrating all units.

The transformation matrices used for the alignment of last disparity image to first disparity image for unit 0, 1, 2 become $\begin{pmatrix} 1.01 & -2.52 \times 10^{-3} & 1.56 \times 10 \\ 1.04 \times 10^{-2} & 9.81 \times 10^{-1} & 2.05 \times 10 \\ 1.56 \times 10^{-5} & 2.23 \times 10^{-6} & 9.86 \times 10^{-1} \end{pmatrix}$, $\begin{pmatrix} 9.94 \times 10^{-1} & -6.04 \times 10^{-4} & -1.02 \times 10 \\ -1.21 \times 10^{-2} & 1.02 & 3.01 \times 10 \\ -1.08 \times 10^{-5} & -1.04 \times 10^{-6} & 1.01 \end{pmatrix}$, $\begin{pmatrix} 9.85 \times 10^{-1} & -4.16 \times 10^{-3} & -1.14 \\ -2.09 \times 10^{-2} & 1.01 & 2.22 \times 10 \\ -1.83 \times 10^{-5} & -2.01 \times 10^{-6} & 1.03 \end{pmatrix}$.

b. Alignment using rotation angles and swing radius with adjustment and optimization:

Figure 7 shows the final height images generated by integrating all units.

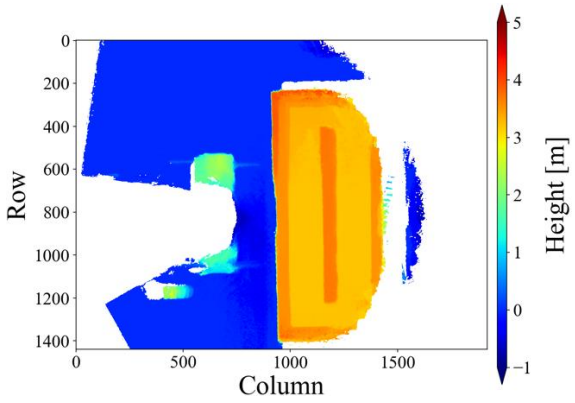

Figure 7: Final height images generated by integrating all units.

The sum of the rotation angles calculated from homography matrices for unit 0, 1, 2 become 1.07×10^{-1}, 9.46×10^{-3}, 2.69×10^{-2} [rad]. In this experiment, all frames used in the outward trip and return trips were the same; thus, all frames can be treated as fixed points. This corresponds to $k = j + 1, u = t + 1$ in Equations (15) and (16). Thus, there are 20 fixed points in each unit, and the rotation differences between the fixed points for unit 0, 1, 2 become as Table 1, 2, 3, where the adjusted angles are colored red.

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

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

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

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

The means for the rotation differences between the fixed points for unit 0, 1, 2 become 5.61×10^{-3}, 4.98×10^{-4} and 1.41×10^{-3} [rad], and the standard deviations for unit 0, 1, 2 become 1.20×10^{-3}, 1.38×10^{-2} and 1.89×10^{-2} [rad]. Thus, according to Equation (16), no rotation angle in unit 0, 2 rotation angles in unit 1, 1 rotation angle in unit 2 are adjusted. After the adjustment, the sums of the rotation angles for unit 0, 1, 2 become 1.07×10^{-1}, 7.45×10^{-3}, 8.07×10^{-2} [rad]. Therefore, each rotation angle in each unit are adjusted -2.81×10^{-3}, -1.96×10^{-4}, -2.12×10^{-3} [rad] because $\varphi_{1n} = 0$ [rad] according to Equation (14), and the sum of rotation angles become -6.10×10^{-6}, -1.65×10^{-3}, 1.59×10^{-3} [rad]. By taking the mean of all swing radius calculated from homography matrices, the integrated swing radius becomes $r = 4.64 \times 10^{2}$ [pix]. Figure 8 shows an example of differences between height images of unit 0 with and without optimization of parameters, which are generated using images in return trip.

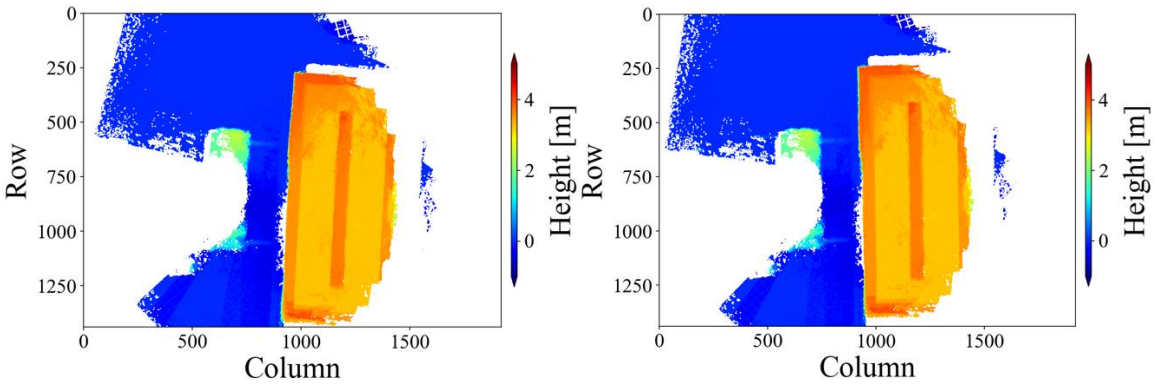

Figure 8: Height images generated by integrating images in return trip in unit 0. Left is whitout optimization and right is with optimization.

Here, we have shown height images using only return trip because the effect of accumulation errors in the alignments is more significant in the latter part of the moving images. Furthermore, in order to see the result of optimization of rotation angles, we used the same radius in the height image without optimization.

Discussion

In this paper, we proposed a new alignment method using parameters such as rotation angles between images and swing radius, an optimization of parameters by a least squared method using calculated rotation angles and swing radius from homography matrices and an adjustment of rotation angles using the differences between the outward trip and return trip of the crane orbit.

a. Alignment using rotation angles and swing radius:

Comparing the result of the alignment using rotation angles and swing radius with that using homography matrices as shown in Figure 6 and 7, it can be said that similar result can be obtained by applying the alignment method not using homography matrices directly but using parameters calculated from homography matrices. We are thinking this is possible because of the feature of the crane orbit, which are a fixed rotation center point of the crane boom without moving of the crane base and a fixed length of the crane boom. However, the error of the center point of the last disparity image is different when we compare them in detail. In this experiment, all frames used in the return trip are completely same with the outward trip. So that, the true transformation matrix for the alignment from the last disparity image to the first disparity image will be an identity matrix. However, when focusing on the translation elements of the result of the alignment using homography matrices, the final disparity image in unit 0, 1, 2 have a translation of about 2.58×10, 3.18×10, 2.22×10 [pix] compared to the first disparity images. Compared to this, since the sum of the rotation angles and swing radius of unit 0, 1, 2 are -6.10×10^{-6}, -1.65×10^{-3}, 1.59×10^{-3} [rad] and 4.64×10^{2} [pix], the final disparity image in unit 0, 1, 2, have a translation of about 2.83×10^{-3}, 7.66×10^{-1}, 7.38×10^{-1} [pix] according to the result of the alignment using rotation angles and swing radius. In the moving images, the camera height h is 15.243 [m] and the focal length of the camera is 931.205 [pix], and 1 [pix] on the final height image's plane equal to 1.61×10^{-2} [m] as real length. So that, it can be said that the accumulation error of the alignment in unit 0, 1, 2 improved from $\times 10^{2}$, 5.12×10^{2}, 3.57×10^{2} [cm] to 4.56×10^{-3}, 1.23, 1.19 [cm] by alignment using rotation angles and swing radius with adjustment and optimization. We are considering that the alignment using parameters is effective because of these constraints of the crane orbit.

b. Optimization of parameters:

Since all frames were the same in the outward and return trips, the sum of the rotation angles of all images should be 0 [rad], so we compare 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 sum of the rotation angles and swing radius before the optimization of unit 0, 1, 2, are 1.07×10^{-1}, 7.45×10^{-2}, 8.07×10^{-2} [rad] and 464 [pix], the final disparity image in unit 0, 1, 2, have a translation of about 4.96×10, 3.46, 3.74×10 [pix]. Those after the optimization of unit 0, 1, 2 are -6.10×10^{-6}, -1.65×10^{-2}, 1.59×10^{-3} [rad] and 464 [pix], the final disparity image in unit 0, 1, 2, have a translation of about 2.83×10^{-3}, 7.66×10^{-1}, 7.38×10^{-1} [pix]. Figure 8 are height images with and without optimization of parameters using images in return trip of unit 0. We only showed the image of unit 0 because the change is larger than in other units. You can see that the rotation error in the left image is adjusted as the right image by the optimization. Thus, you can see that the accumulation errors of the sum of the rotation angles and the estimated location of the last disparity image improved. However, there are also other possible residuals for the optimization method. For example, a residual using distances between principal points $\varepsilon_i = b_i - r\theta_i$ can be used in the least squares method. In this paper, we did not use this residual because the accuracy of the distances between principal points depends on other parameters like a scale adjustment factor.

c. Adjustment of parameters:

When we focus on the result just after the adjustment using fixed points, the sum of the rotation angles of unit 0, 1, 2 change from 1.07×10^{-1}, 9.46×10^{-3}, 2.69×10^{-2} [rad] to 1.07×10^{-1}, 7.45×10^{-2}, 8.07×10^{-2} [rad], and the sum differs from the true sum 0 in unit 2. Considering that there is not much difference in the means and standard deviations for the rotation differences between the fixed points in units, the adjustment limited to the return trip might be the reason. In this method, we adjust the rotation angle of the return trip to that of the outward trip; however, it is impossible to judge which trip's rotation angle includes errors. Thus, there is a possibility that the rotation angle of the outward trip is incorrect. For the adjustment, it might be necessary to consider which rotation angle to adjust by comparing the rotation angle between other fixed points.

Conclusion

In order to improve the labor productivity at construction sites, an automatic crane operation is attracting attention, and a three-dimensional map of a crane surroundings is necessary for the crane system to recognize the surroundings and to determine safe lifting

routes of the loads. For a generation of three-dimensional map, a method using disparity images generated from moving images taken from a tip of the crane boom using monocular have been developed, but this method has an accumulation error of the alignment of each disparity image because it repeatedly utilizes homography matrices calculated from each pair images for the alignment. In order to deal with this problem, we proposed a new alignment method using rotation angles and swing radius with processes of adjustment and optimization of the parameters. For the adjustment, we applied the differences of rotation angles between fixed points in outward trip and return trip, and for the optimization, we used least square method using residuals as rotation angles and swing radius calculated from homography matrices. By using the proposed alignment method, the error from the true location of the last disparity image have been improved compared to the previous alignment error. So that, despite the limited situation of using the same image for the outward and return trips, it can be said that we were able to propose a method to adjust the accumulation error in the alignment of the disparity images. In the future we aim to validate our current implementation on in-situ data of a real crane round trip motion, and extend the method to account for changing 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, November 2023.*

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*