

# Visualization of Dump Truck and Excavator in Bird's-eye View by Fisheye Cameras and 3D Range Sensor

Yuta Sugasawa<sup>1</sup>, Shota Chikushi<sup>1</sup>, Ren Komatsu<sup>1</sup>,  
Jun Younes Louhi Kasahara<sup>1</sup>, Sarthak Pathak<sup>1</sup>, Ryosuke Yajima<sup>1</sup>,  
Shunsuke Hamasaki<sup>1</sup>, Keiji Nagatani<sup>1</sup>, Takumi Chiba<sup>2</sup>, Kazuhiro Chayama<sup>2</sup>,  
Atsushi Yamashita<sup>1</sup>, and Hajime Asama<sup>1</sup>

<sup>1</sup> The University of Tokyo

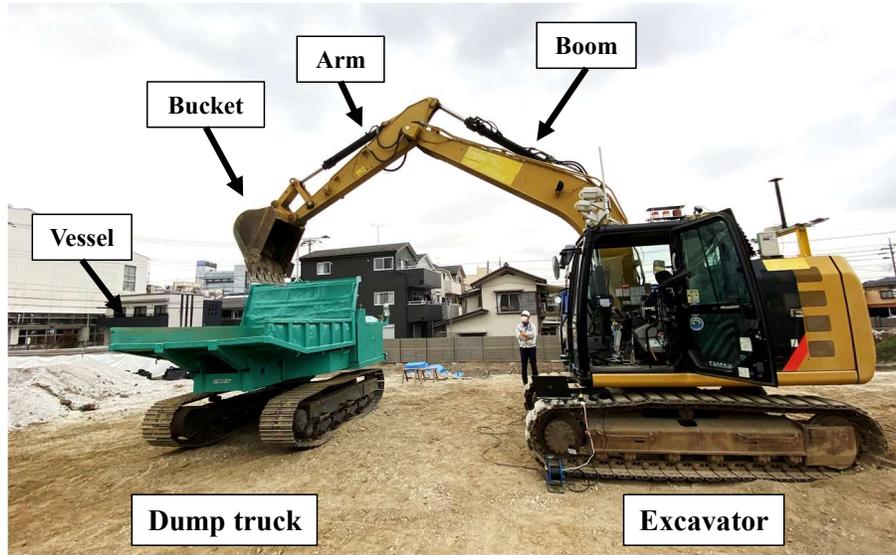
<sup>2</sup> Fujita Corporation

**Abstract.** In disaster sites, workers may not be allowed entry due to the risk of secondary disasters. Remote operation of construction machines is then necessary for emergency disaster recovery work. For remote operation of construction machines, a third-person view is effective. External cameras, such as mobile camera vehicles or fixed cameras, are commonly used to display the third-person view in remote construction sites. However, if mobile camera vehicles or external cameras can not be installed on-site, displaying a third-person view is not possible. To solve this problem, in this paper, using a 3D range sensor and fisheye cameras mounted onboard the remotely operated construction machine, we measure the positions and poses of other construction machines and display their 3D models inside a bird's-eye view which is a reconstructed 3D view. This allows the remote operator to recognize the relative position and pose and easily operate the machine.

**Keywords:** Bird's-eye View, Fisheye Camera, Remote Operation, Construction Machine

## 1 Introduction

In the event of a disaster, remote operation of construction machines is necessary for emergency disaster recovery [1]. However, there are some disaster sites where it is difficult for construction machine operators to enter due to the risk of secondary disasters. For example, at a disaster site where a landslide has occurred, there may be the risk that another landslide will occur unexpectedly. For such disaster sites, remote operation of construction machines is effective. In remote operation of construction machines, construction equipment is installed at the disaster sites and remotely operated by remote operators to conduct recovery work. One of the most common tasks in such scenarios is earth-moving work using excavators and dump trucks. The earth-moving work is carried out by the operator of the excavator at a remote location while viewing the images from the on-board cameras and the external cameras installed at the construction site.



**Fig. 1.** Earth-moving work with excavator and dump truck

A mobile camera vehicle equipped with cameras that can change its viewpoint is commonly used as the external camera, but a fixed camera installed at the construction site may also be used. When neither a mobile camera vehicle nor a fixed camera can be used, wired-powered drones may be used in actual work sites. A wired-powered drone receives electricity from an excavator and provides images to the remote operator from above the excavator.

Figure 1 shows the earth-moving work with the excavator and the dump truck. As shown in Fig. 1, the excavator has the following components: boom, arm, and bucket. These parts rotate based on their joints, allowing the excavator to do earth-moving work. Sand excavated by the excavator is loaded onto the vessel of the dump truck. Therefore, recognizing the relative positions of the excavator and the dump truck is important for earth-moving work.

Remote operation is commonly performed while viewing multiple third-person views provided by external cameras. Multiple third-person views allow the remote operator to change viewpoints depending on the task, and it is easy to recognize the relative positions of the excavator and the dump truck. For example, when excavating soil, the remote operator has to look at the relative position of the bucket of the excavator and soil. For another example, when loading soil, the remote operator has to look at the relative position of the vessel of the dump truck and the bucket of the excavator. Therefore, multiple third-person views are useful for remote operation of earth-moving work [2][3].

However, in disaster sites, the installation of external cameras requires a lot of time. For example, if the environment of the construction site is not maintained, mobile camera vehicles may not enter the sites, and fixed cameras may

not be installed. Wired-powered drones may also not be used in bad weather conditions such as rain and wind. Therefore, for the remote earth-moving work, this research focuses on the display of a third-person view without using an external camera. This third-person view allows remote operators to smoothly manipulate excavators.

There are some previous researches that displayed third-person views without using external cameras. Sun et al. installed four fisheye cameras on a crawler-type mobile robot and integrated the images from each camera to generate a bird's-eye view which was projected on a hemispherical dome centered on the crawler-type mobile robot [4]. However, because the images from the fisheye cameras were projected on the hemispherical dome, the position, pose, and shape of tall objects around the crawler-type mobile robot could not be displayed correctly. In the remote operation of the excavator and the dump truck, this problem is fatal because it does not allow the display of construction machines performing cooperative work.

Komatsu et al. proposed a method for displaying a wall in an indoor environment by using a mobile robot equipped with four fisheye cameras and a laser range finder [5]. Under the assumption that the wall is perpendicular to the floor, they displayed the position and shape of the wall in a bird's-eye view. However, that assumption is too simplistic to be effectively applicable to construction machines.

Based on the above, none of the previous researches can be applied to an excavator remotely operated for earth-moving work. Therefore, the purpose of this research is to generate a third-person view from onboard cameras mounted on an excavator without using external cameras.

## 2 Proposed method

### 2.1 Overview

First of all, we define the bird's-eye view in this research. The bird's-eye view is the view that plays the role of multiple third-person views, and the user can arbitrarily change the viewpoint. The bird's-eye view is generated by projecting the four images of the four fisheye cameras mounted on an excavator onto a hemispherical dome. The method of generating bird's-eye view is based on the method of Komatsu et al. [5]. The remote operator can see the excavator from a third-person's point of view by using the bird's-eye view, and the viewpoint can be changed depending on the task. Therefore, in the scenario of remote operation without external cameras, the bird's-eye view generated from the four fisheye cameras is an alternative to multiple third-person views and is a suitable image presentation method for the remote operator of the excavator.

The novelty of this research is that it proposes a new method for displaying a position, pose and shape of an excavator and a dump truck based on estimated position and pose of the dump truck in a bird's-eye view. This research uses a 3D range sensor for estimating the position and pose of a dump truck and four

fish-eye cameras for generating a bird’s-eye view. These sensors are mounted on an excavator.

In disaster sites, GNSS signals are blocked by surrounding objects such as trees, mountains, etc. Therefore, a 3D range sensor is used in this research, considering that GNSS may not be sufficiently available.

Three-dimensional models of the excavator and the dump truck are used to display the position, pose, and shape of the excavator and the dump truck in the bird’s-eye view. The 3D models are created in advance by measuring the shapes of the excavator and dump truck. The 3D models of the excavator and the dump truck are displayed in the bird’s-eye view based on the estimated positions and poses.

The proposed method can be divided into two steps. The first step estimates the position and pose of the excavator and the dump truck. The second step generates a bird’s-eye view from four fish-eye cameras mounted on the excavator based on the method of Komatsu et al. [5], and displays the 3D models of the excavator and the dump truck in the bird’s-eye view based on the estimated positions and poses.

## 2.2 Definition of the position and pose in three coordinate systems

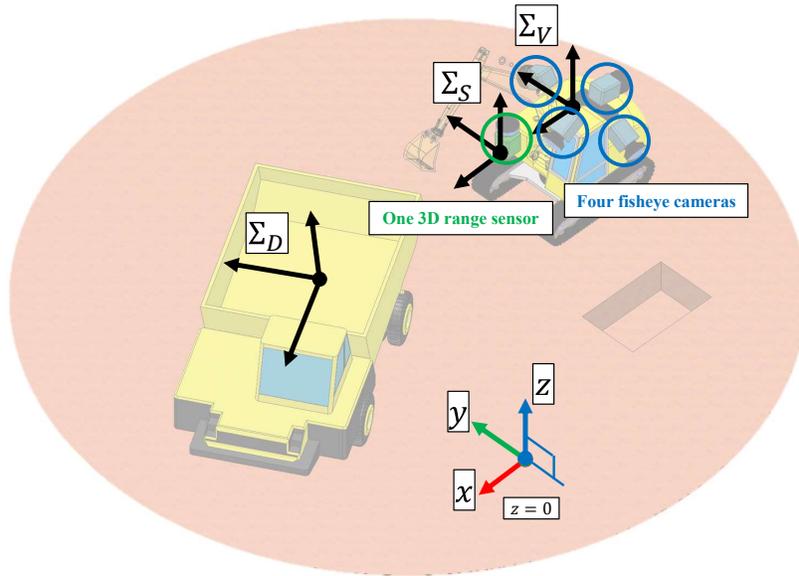
Figure 2 shows the overview of the proposed method. In this research is considered a scenario with one excavator and one dump truck, with no additional construction machine in the surroundings. The excavator and the dump truck are remotely operated and conduct earth-moving work. Furthermore, the bird’s-eye view is supposed to be displayed to the remote operator of the excavator.

In the considered scenario, there are three coordinate systems, as illustrated in Fig. 2. These coordinate systems are explained in Table 1. In this research, the excavator and the dump truck are located on the same plane, and the coordinate of that plane is set to  $z = 0$ . In all three coordinate systems, the  $z$ -axis is oriented perpendicularly to the plane. Therefore, the positions of the excavator and the dump truck are represented by only their  $x$  and  $y$  coordinate values. The pose is represented by only the rotation angle with respect to the  $z$ -axis.

In the remainder of this paper, the position of an object named  $b$  in the coordinate system  $\Sigma_X$  is denoted as  $t_{b,X}$ , its pose as  $R_{b,X}$ . For example, the position of the excavator in the bird’s-eye view’s coordinate system  $\Sigma_V$  is  $t_{\text{excavator},V}$ , and its pose is  $R_{\text{excavator},V}$ .

As shown in Fig. 2, four fish-eye cameras are mounted on the excavator to generate the bird’s-eye view. One 3D range sensor is equipped on the excavator to estimate the position and pose of the dump truck in the 3D range sensor’s coordinate system  $\Sigma_S$ .

The 3D models of the excavator and the dump truck are prepared in advance based on actual measurements. These 3D models are displayed in the bird’s-eye view. The positions and poses of the four fish-eye cameras and one 3D range sensor with respect to the excavator are obtained in advance using the 3D model of the excavator.



**Fig. 2.** Considered scenario with one excavator and one dump truck

**Table 1.** Three coordinate systems in Fig. 2

Coordinate system	Explanation
$\Sigma_V$	The bird's-eye view's coordinate system
$\Sigma_S$	The 3D range sensor's coordinate system
$\Sigma_D$	The dump truck's coordinate system

### 2.3 Estimation of the position, pose, and joint poses of the excavator

To display the 3D models of the excavator and the dump truck in a bird's-eye view, it is necessary to estimate the position and pose of the excavator and the dump truck. When generating the bird's-eye view from four fisheye cameras based on the method of Komatsu et al. [5], the positions and poses of the four fisheye cameras mounted on the excavator in the bird's-eye view's coordinate system  $\Sigma_V$  are obtained by calibration using AprilTag [6][7].

The position and pose of the excavator in the bird's-eye view's coordinate system  $\Sigma_V$  are estimated from the positions and poses of the four fisheye cameras. This is because the position and pose of the four fisheye cameras in relation to the excavator are known from the 3D model of the excavator that was created beforehand based on the actual measurements.

In addition, tiltmeters are mounted on each joint of the arm, boom, and bucket of the excavator to obtain the angle information of each joint. The angle information provides joint poses that refer to the pose of each joint of the excavator. Using the above methods, the position and pose in the bird's-eye view's coordinate system  $\Sigma_V$ , and the joint poses of the excavator can be measured.

#### 2.4 Estimation of the position and pose of the dump truck

The position and pose of the dump truck in the 3D range sensor's coordinate system  $\Sigma_S$  can be estimated using a 3D range sensor mounted on the excavator. Iterative Closest Point (ICP) [8] is used to match the point cloud of the 3D model of the dump truck prepared in advance with the point cloud of the dump truck measured by the 3D range sensor. Matching a point cloud means to move the point cloud so that the distance between two point clouds is minimized.

By matching the point cloud of the dump truck measured by the 3D range sensor with the point cloud of the dump truck prepared in advance, the transformation matrix  $T_{\text{dump}}$  can be obtained.  $T_{\text{dump}}$  contains the position  $t_{\text{dump},S}$  and the pose  $R_{\text{dump},S}$  in the 3D range sensor's coordinate system  $\Sigma_S$ .  $T_{\text{dump}}$  is shown as in (1).

$$T_{\text{dump}} = \begin{bmatrix} R_{\text{dump},S} & t_{\text{dump},S} \\ 0 & 1 \end{bmatrix}. \quad (1)$$

$T_{\text{dump}}$  represents the position and pose of the dump truck in the 3D range sensor's coordinate system  $\Sigma_S$ , but to display the 3D model of the dump truck in the bird's-eye view, the position  $t_{\text{dump},V}$  and the pose  $R_{\text{dump},V}$  in the bird's-eye view's coordinate system  $\Sigma_V$  are needed. Therefore,  $T_{\text{sensor}}$  that represents the position and the pose of the 3D range sensor in the bird's-eye view's coordinate system  $\Sigma_V$  is needed.

Because 3D model of the excavator is prepared in advance, the position and pose of the 3D range sensor relative to the excavator can be obtained. Therefore,  $T_{\text{sensor}}$  is obtained by using the position and pose of the excavator in the bird's-eye view's coordinate system  $\Sigma_V$ .

According to the above,  $S_{\text{dump}}$ , representing the position and pose of the dump truck in the bird's-eye view's coordinate system  $\Sigma_V$ , is calculated as in (2).

$$S_{\text{dump}} = T_{\text{sensor}} T_{\text{dump}} = \begin{bmatrix} R_{\text{dump},V} & t_{\text{dump},V} \\ 0 & 1 \end{bmatrix}. \quad (2)$$

#### 2.5 Display of 3D models in bird's-eye view

Based on the estimation of the position and pose of the excavator and dump truck, the 3D models of the excavator and dump truck are displayed in the bird's-eye view. This allows the remote operator to recognize the relative positions and poses of the excavator and the dump truck in the bird's-eye view, enabling smooth remote operation of the excavator for earth-moving work.

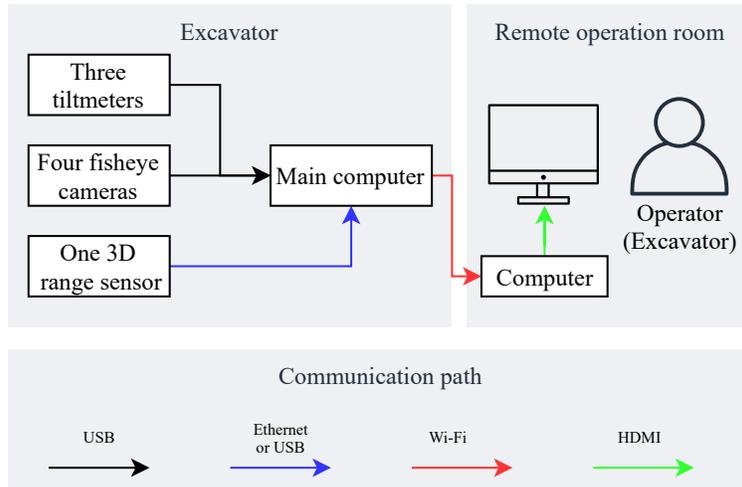


Fig. 3. System architecture diagram

## 2.6 System architecture diagram of the proposed method

Figure 3 shows the system architecture diagram of the proposed method. The three tiltmeters acquire the rotation angles of the joints of the boom, arm and bucket of the excavator. Four fisheye cameras acquire 360-degree images of the excavator's surroundings. One 3D range sensor acquires the point cloud of the dump truck.

The data from the above sensors are transferred to the main computer installed on the excavator, and the bird's-eye view where 3D models of the excavator and the dump truck are displayed is generated. The bird's-eye view is transferred wirelessly to the computer installed in the remote operation room. The operator can remotely control the excavator by checking the bird's-eye view on the monitor installed in the remote operation room.

## 3 Experiment

### 3.1 Experimental setting

A field experiment was conducted with the excavator and the dump truck. The experimental setting is shown in Fig. 4. VLP-16 developed by Velodyne Lidar was used as the 3D range sensor and mounted on the excavator. The VLP-16 emits 16 line lasers in all directions and acquires the point cloud of the dump truck. Among the point cloud of the dump truck acquired by VLP-16, the 2D point cloud that best represented the features of the dump truck shape was extracted manually. That 2D point cloud of the dump truck is used for point cloud matching by ICP.

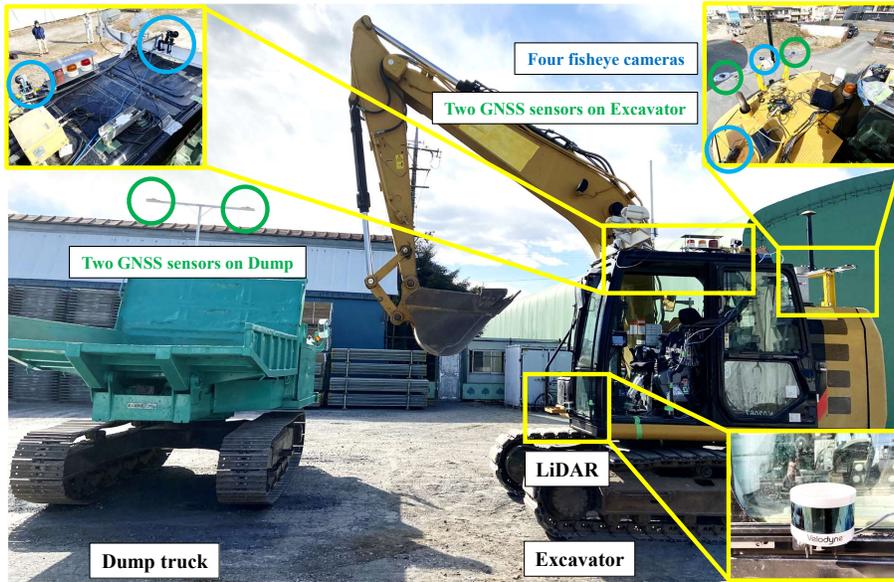


Fig. 4. Experimental setting

The horizontal angular resolution of the VLP-16 is 0.1 - 0.4 [deg]. The distance between the excavator where the VLP-16 was mounted and the dump truck was at most 5 [m], which was the working range of the excavator. In this case, the maximum horizontal measurement error of the point cloud of the dump truck obtained by VLP-16 is about 0.03 [m]. Considering the total length of the dump truck is about 5 [m], the measurement error is not a problem in point cloud matching.

The 3D models of the excavator and the dump truck displayed in the bird's-eye view were obtained by Structure from Motion (SfM). Each of these 3D models was scaled to match the actual size of the machines. By matching the point cloud obtained by cutting out a part of the 3D model of the dump truck obtained by SfM with the point cloud obtained by LiDAR, the position and pose of the dump truck  $T_{\text{dump}}$  in the 3D range sensor's coordinate system  $\Sigma_S$  was estimated.

In this experiment,  $T_{\text{sensor}}$ , which is the position and pose of LiDAR mounted on the excavator in the bird's-eye view's coordinate system  $\Sigma_V$ , is calculated by manually selecting one point for the sensor port of LiDAR from the point cloud of the excavator obtained by SfM. According to the official manual of VLP-16, the LiDAR has a cylindrical shape with a diameter of 0.103 [m], and a point in the center of the cylindrical shape was manually selected. Because one point was selected manually from multiple points,  $T_{\text{sensor}}$  may have carried an error of at most 0.103 [m].

In Fig. 4 are shown the four fisheye cameras on the top of the excavator, which are indicated by the blue border in the figure. Furthermore, in order to

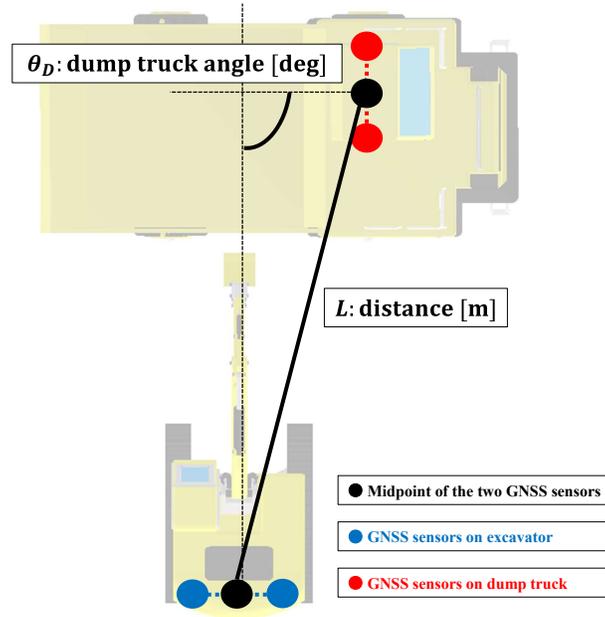


Fig. 5. Top view of the experimental setting

obtain the true value of the relative position and pose of the excavator and the dump truck, two GNSS sensors were mounted on the top of both the excavator and the dump truck. The GNSS sensors are indicated by the green border in the figure. These GNSS sensors received correction signals from the neighboring base stations, and are capable of Real-time kinematic (RTK) processing. The error of the RTK-GNSS sensor is in the order of a few centimeters.

In this experiment, the position and pose of the dump truck relative to the excavator calculated by the proposed method are compared with the true values calculated by GNSS. Fig. 5 shows a view of the experimental settings from above. The blue circles in Fig. 5 show the two GNSS sensors on the excavator and the red circles show the two GNSS sensors on the dump truck. The black circles show the midpoint of the two GNSS sensors on each construction machine. In Fig. 5, the distance between the two midpoints is distance  $L$ , and the relative angle of the dump truck to the excavator is dump truck angle  $\theta_D$ .

The compared values are distance  $L$  and dump truck angle  $\theta_D$ . The unit of distance  $L$  is [m], and the unit of dump truck angle  $\theta_D$  is [deg]. Both distance and dump truck angle are expressed as values for the bird's-eye view coordinate system  $\Sigma_V$ .



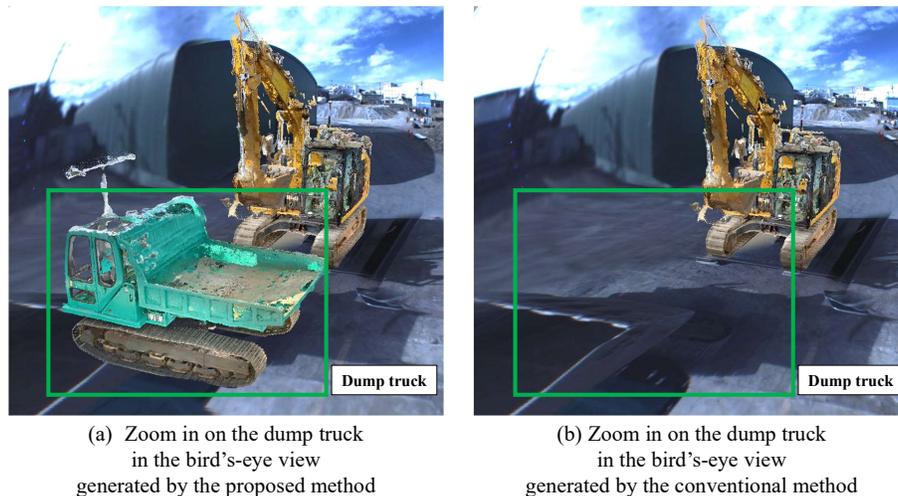
**Fig. 6.** Comparison of the bird's-eye views generated by the proposed method and the conventional method [5].

### 3.2 Experimental result

The error between the calculated value of the distance  $L$  and dump truck angle  $\theta_D$  by the proposed method and the true value by GNSS was obtained. The error of the distance  $L$  was 0.310 [m]. The error of the dump truck angle  $\theta_D$  was 2.405 [deg].

Because RTK-GNSS is accurate to within a few centimeters, the error is considered to have been compounded in the work of the proposed method. Specifically, when scaling the point clouds of the excavator and the dump truck obtained by SfM to actual size, and when selecting one point of LiDAR in the point cloud of the excavator to calculate  $T_{\text{sensor}}$ , the work was done manually, and thus there is a possibility of errors due to manual work. The error is also expected to ride when  $T_{\text{dump}}$  is calculated by ICP. In this experiment, it is not possible to evaluate which error is dominant. It is also difficult to quantitatively evaluate how much the error of distance  $L = 0.301$  [m] affects the earth-loading work because the remote operator does not actually operate the truck remotely in our experiments.

Figure 6 shows the bird's-eye view generated by the proposed method. The hemispherical dome of the bird's-eye view has a radius of 10.0 [m]. In Fig. 6(a)



**Fig. 7.** Comparison of the two views taken from different angles generated by the proposed method and the conventional method [5].

is shown a photo taken by a perspective camera of the experimental conditions. Fig. 6(b) shows the entire bird's-eye view, and the viewpoint in this bird's-eye view can be changed arbitrarily. Fig. 6(c) is the enlarged view of the yellow frame in Fig. 6(b), and the green frame in Fig. 6(c) and Fig. 6(d) shows the position of the dump truck. Fig. 6(d) is the bird's-eye view generated by the conventional method [5].

As shown in Fig. 6, Fig. 6(b) and Fig. 6(c) are similar to Fig. 6(a). In addition, when comparing Fig. 6(c) and Fig. 6(d), the position, pose and shape of the dump truck is displayed correctly in Fig. 6(c). However, in Fig. 6(d), the relative position and pose of the excavator and the dump truck in the bird's-eye view can not be recognized.

Figure 7 shows the bird's-eye views cut out from a different angle than Fig. 6. Fig. 7(a) is generated by the proposed method, and Fig. 7(b) is generated by the conventional method [5]. In Fig. 7(a), the dump truck surrounded by a green frame is displayed as a 3D model, but in Fig. 7(b), it is not. Therefore, in Fig. 7, the remote operator can not recognize the relative positions of the bucket of the excavator and the vessel of the dump truck. On the other hand, in Fig. 7(a), the relative positions are visible.

## 4 Conclusion

In this research, we proposed a method to display the 3D models of the excavator and the dump truck in the bird's-eye view based on the estimation of their positions and poses. In the experiment, we confirmed that the proposed method was effective because the 3D models of the excavator and the dump truck were

displayed in the bird's-eye view. By using the proposed method, it is possible to generate the bird's-eye view from four fisheye cameras mounted on the excavator and display the position, pose, and shape of the excavator and dump truck in the bird's-eye view even in an environment where external cameras can not be installed. The remote operator of the excavator can recognize the position, pose, and shape of the excavator and dump truck by checking the displayed bird's-eye view, and it is expected to enable smooth operation of the earth-moving work.

As future work, we would like to pursue development of our proposed method to operate in real-time. For this aim, it will be necessary to develop a method for estimating the position and pose of the dump truck for real-time operation. In addition, it will be necessary to consider a method to quantitatively compare the position, pose and shape of an excavator and a dump truck when they are displayed in a bird's-eye view, along with how faithfully they display the position, pose, and shape of the actual excavator and dump truck. Following that, we would like to conduct real-time field experiments.

## References

1. Chayama K., Fujioka A., Kawashima K., Yamamoto H., Nitta Y., Ueki C., Yamashita A. and Asama H.: Technology of Unmanned Construction System in Japan. *Journal of Robotics and Mechatronics*, Vol. 26, No. 4, pp. 403-417, 2014.
2. Sato R., Kamezaki M., Sugano S. and Iwata H.: Gaze Pattern Analysis in Multi-display Systems for Teleoperated Disaster Response Robots. *Proceedings of the 2016 IEEE International Conference on Systems, Man and Cybernetics*, pp. 3534-3539, 2016.
3. Kamezaki M., Yang J., Iwata H. and Sugano S.: Visibility Enhancement Using Autonomous Multicamera Controls with Situational Role Assignment for Teleoperated Work Machines. *Journal of Field Robotics*, Vol. 33, No. 6, pp. 802-824, 2016.
4. Sun W., Iwataki S., Komatsu R., Fujii H., Yamashita A. and Asama H.: Simultaneous Tele-visualization of Construction Machine and Environment Using Body Mounted Cameras. *Proceedings of the 2016 IEEE International Conference on Robotics and Biomimetics*, pp. 382-387, 2016.
5. Komatsu R., Fujii H., Tamura Y., Yamashita A. and Asama H.: Free Viewpoint Image Generation System Using Fisheye Cameras and a Laser Rangefinder for Indoor Robot Teleoperation. *ROBOMECH Journal*, Vol. 7, 15, pp. 1-10, 2020.
6. Olson E.: AprilTag: A Robust and Flexible Visual Fiducial System. *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*, pp. 3400-3407, 2011.
7. Wang J. and Olson E.: AprilTag 2: Efficient and Robust Fiducial Detection. *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4193-4198, 2016.
8. Rusinkiewicz S. and Levoy M.: Efficient Variants of the ICP Algorithm. *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145-152, 2001.