A Study on Autonomous Operation System of Caisson Shovels in High Air Pressure and Narrow Underground Space

(Terrain Mapping by RGB- and Depth-sensing Camera Mounted on the Shovel)

Akira KAMEI*[1], Koki KIKUCHI*[2], Tetsuya KOYO*[1] and Toshihiro KONDO*[1]

*1 Oriental Shiraishi Corp.
5-6-52 Toyosu, Koto-ku, Tokyo 135-0061, JAPAN
Akira.kamei@orsc.co.jp

*2 The Department of Advance Robotics, Chiba Institute of Technology
2-17-1 Tsudanuma, Narashino, Chiba 275-0016, JAPAN
kikut@ieee.or.jp

Abstract

In this paper, we focused on unmanned construction by autonomous excavators in a high air pressure and narrow underground space and developed a 3D terrain mapping system using a depth sensor. This system consisted of a shovel designed as a linkage mechanism with five degrees of freedom and an RGB- and depth-sensor mounted on the shovel and mapped by translating depth cloud points from the camera coordinate system into digital elevation model (DEM) information. In order to evaluate the accuracy and reliability of this system, we mapped the artificial terrain that arranged rectangular columns with heights of 100mm to 500mm on a flat plate for two cases: Case 1 maps based on an image from a camera on a fixed shovel (single view) and Case 2 maps based on images from a camera on a shovel traveling on the rail track (time series view). The results revealed that the error in Case 2 was a few centimeters or less and the accuracy and reliability of the terrain map integrated from multiple information sources was sufficiently high for the excavation operation.

Keywords: unmanned construction, excavation, terrain mapping, caisson shovel

1 Introduction

The demand for large underground structures, such as bridge piers, pump stations for sewage treatment facilities, water storage tanks, and vertical shafts for subways and tunnels, is high. A pneumatic caisson method (PCM) [1], [2], a typical excavation method, constructs a reinforced concrete caisson box with a working chamber, as shown in **Fig. 1**, excavates the ground by teleoperated shovels mounted on the ceiling of the working chamber, and immerses vertically the caisson box into the ground gradually. The feature is to excavate the ground while avoiding groundwater intrusion by high air pressure corresponding to the groundwater pressure. Meanwhile, recently several issues and problems have been identified. For example, skilled operators are aging and new entry operators tend to lack skill and safety consciousness. This leads to accidents such as shovel collisions and trivial machine troubles. Such an absolute shortage in the workforce will lead to a gradual decrease in the construction efficiency. From this reason, we focus on an unmanned construction, i.e., the automation of excavation by autonomous caisson shovels, which replaces all or a part of excavation works by human operators. Here, autonomous caisson shovels in the working chamber have to recognize the three dimensional roughness of the ground for excavation and plan the excavation sequence automatically, since the average height of the working chamber keeps 2.3m to immerse the caisson box vertically and to allow the bucket of the shovel to reach the ground. In addition, the whole 3D roughness of the ground in the working chamber must be updated in real time (synchronously with the operation) as the terrain map.

In the construction industry, 3D mapping using an unmanned aerial vehicle (UAV) has started for terrain measurement [3]-[8]. These technologies obtaining the terrain information by the combination among global positioning system (GPS), a laser range finder, and aerial images created a wide range of terrain map with an accuracy of cm-order. In a PCM, however, since the working chamber is narrow (ceiling height: approximately 2.3m) and sealed, as shown in **Fig. 2**, a UAV system is difficult to use and a global positioning system (GPS) cannot be used. Meanwhile, another type of system has been used for mapping in actual earthwork sites for the purpose of efficient construction [9], [10]. This system created a 3D map integrating the information obtained from multiple laser scanners with a visual range of 360deg, which were located at many fixed observation points. This map is updated, for example, only once a week because of high renewal cost. But the map for the excavation planning must be updated synchronously during the excavation operation. Yamamoto *et al*. [11], [12] proposed a 3D terrain measurement method for autonomous excavation using a stereo camera and a laser scanner mounted on a backhoe. Although the vision information from the front of the backhoe mapped the excavation area partially, excavation planning in PCM requires a whole terrain map to immerse the caisson box vertically and precisely.

From these points of view, we propose a 3D terrain mapping method integrating information obtained from

multiple RGB- and Depth-sensing (RGBD) cameras mounted on the caisson shovels. The RGBD sensor obtains depth information for the geometric mapping and an RGB image for soil property estimation simultaneously. In addition, while an individual RGBD sensor recognizes the neighborhood in real time, multiple RGBD sensors create a whole terrain map periodically. In this paper, as a first step in the development of this unmanned construction system, we focus on 3D terrain mapping based on depth information obtained by a single RGBD sensor on a caisson shovel and investigate the accuracy and reliability of the system.

The remainder of this paper is organized as follows. In Section 2, we describe the architecture of the 3D terrain mapping system. In Section 3, we explain the mathematical model of the caisson shovel and RGBD sensor position. In Section 4, we describe the experimental environment used to investigate the accuracy of the mapping system. In Section 5, we present the experimental results and discuss the characteristics of the mapping system. Finally, in Section 6, we conclude the paper and outline future works.

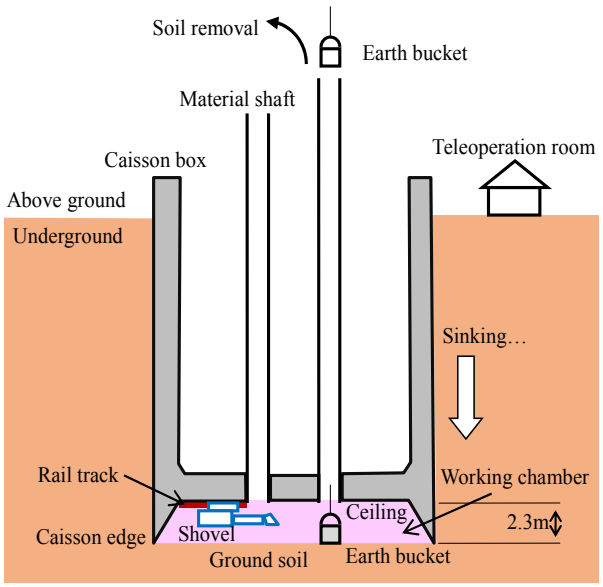

Fig. 1 Pneumatic caisson facility

Fig. 2 Example of ground condition below working chamber

2 Architecture of terrain mapping system

2.1 Caisson shovel

Figure 3 shows the caisson shovel, which is fixed on a rail track on the working chamber ceiling and stably exerts a larger bucket force compared to a common backhoe that receives the reaction force from unstable rough ground. The caisson shovel consists of four parts: carriage, boom, counter weight, and bucket units. These units are inserted through the material shaft and are assembled in the working chamber. After construction is complete, these are disassembled and ejected through the material shaft again. The caisson shovel has five degrees of freedom: traveling of carriage unit on the rail track (d_0), yaw rotation of the boom unit (θ_1), pitch rotation (dumping) of the boom unit (θ_2), expansion and contraction of the boom unit (d_3), and pitch rotation of the bucket unit (θ_4), as described in Subsection 3.1. The caisson shovel is driven by hydraulic cylinders using an external electric power supply. Since each joint has an internal sensor, the position and orientation of the caisson shovel from the viewpoint of the working camber coordinate system are calculated uniquely.

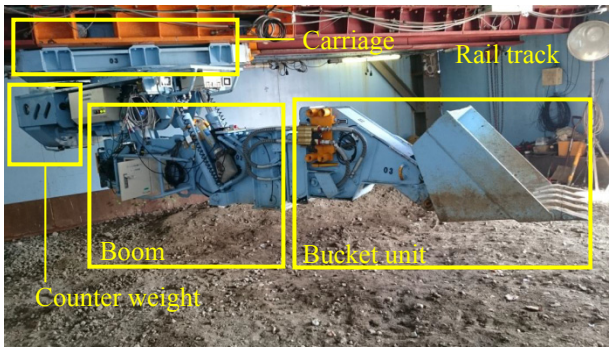

Fig. 3 Caisson shovel

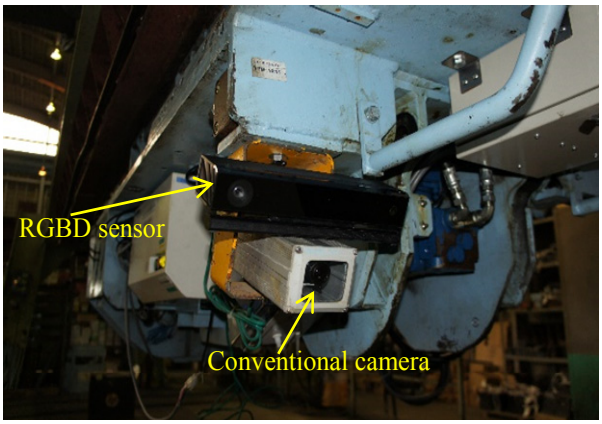

Fig. 4 RGBD sensor mounted on caisson shovel

2.2 RGBD-sensing camera

In this study, we use an RGBD sensor that provides a 1920pixel×1080pixel resolution RGB image and a 512pixel×424pixel resolution depth image as a terrain recognizing sensor. These angular fields of view are 84.1deg×53.8deg for the RGB image and 70.6deg×60.0deg for the depth image, respectively.

Note that although the center of the RGB image does not correspond to that of the depth image, the API software matches these images to the origin of the coordinate system. In addition, the depth sensor range is from 0.5m to 8.0m and the out of range is recognized as 0.0m. The depth resolution is 1mm. As shown in **Fig. 4**, the RGBD sensor is attached to the boom unit as well as the conventional camera for teleoperation, moves according to the traveling of the carriage unit on the rail track (d_0), and rotates according to the yaw angle rotation (θ_1) of the boom unit. Note that, since our target space, i.e., the inside of the working chamber, is always sealed and illuminated by constant brightness, the lighting condition is constant and stable for image processing. However, slight reflections by the underwater are observed.

3 Mathematical model

3.1 Kinematics of the caisson shovel

Figure 5 shows the mathematical link manipulator model for the caisson shovel and the coordinate systems used in the analysis. This is the initial configuration when $d_0 = \theta_1 = \theta_2 = d_3 = \theta_4 = 0$. Here, Σ_R is the frame of reference located at the ceiling of the working chamber and Σ_H is the coordinate system for the bucket claw. **Table 1** shows the link parameters [1] from Σ_1 to Σ_H. From the link parameters and the relationship between Σ_R and Σ_H, the homogeneous transformation matrix is obtained as

$$ {}^{R}T_{H} = \begin{bmatrix} C_1C_{2+4} & -S_1 & -C_1S_{2+4} & p_x \\ S_1C_{2+4} & C_1 & -S_1S_{2+4} & p_y \\ S_{2+4} & 0 & C_{2+4} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

where

$$ \begin{aligned} p_x &= C_1(l_6S_{2+4} + l_5C_{2+4} - l_3C_2 + (d_3 + l_4)S_2 - l_2) + d_0 \\ p_y &= S_1(l_6S_{2+4} + l_5C_{2+4} - l_3C_2 + (d_3 + l_4)S_2 - l_2) \\ p_z &= -l_6C_{2+4} + l_5S_{2+4} - l_3S_2 - (d_3 + l_4)C_2 - l_1 \end{aligned} \quad (1) $$

Here $S_i \equiv \sin\theta_i$ and $C_i \equiv \cos\theta_i$. From this, the relationship between the position and orientation of the caisson shovel claw, $\boldsymbol{r}$, is

$$ \boldsymbol{r} = \begin{bmatrix} x_H \\ y_H \\ z_H \\ \Phi \\ \Theta \\ \Psi \end{bmatrix} = \begin{bmatrix} p_x \\ p_y \\ p_z \\ \theta_1 \\ \theta_2 + \theta_4 \\ 0 \end{bmatrix} \quad (2) $$

where Φ, Θ, and Ψ are the yaw, pitch, and roll angles, respectively. Hence, this mechanism cannot change the roll angle. Except for the role angle, the position and orientation are uniquely determined by five DOF variables. The link lengths are as follows: l_1 = 829 [mm], l_2 = 170 [mm], l_3 = 415 [mm], l_4 = 2,631 [mm], l_5 = 813 [mm], and l_6 = 259 [mm].

Next, assuming that d_0=0 under excavation, we obtain the inverse kinematics of the shovel, since the carriage unit is fixed on the rail track to support the reaction force stably during the actual excavation operation. From the inverse kinematics, the position and pitch angle of the shovel claw are obtained as

$$ \begin{bmatrix} \theta_1 \\ \theta_2 \\ d_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} \arctan_2(y_H, x_H) \\ \arctan_2(k_1, k_2) - \arctan_2(\sqrt{k_1^2 + k_2^2 - l_3^2}, l_3) \\ (x_H / C_1 - l_6S_\Theta - l_5C_\Theta + l_3C_2 + l_2) / S_2 - l_4 \\ \Theta - \theta_2 \end{bmatrix} \quad (3) $$

where

$$ k_1 = -l_6C_\Theta + l_5S_\Theta - l_1 - z_H $$
$$ k_2 = -l_6S_\Theta + l_5C_\Theta - l_2 - y_H / S_1 $$

Table 1 Link parameters

i	a_{i-1}	α_{i-1}	d_i	θ_i
1	*0*	0	$-l_1$	θ_1
2	$-l_2$	90°	0	θ_2
3	$-l_3$	90°	d_3+l_4	0
4	0	-90°	0	θ_4
H	l_5	-90°	$-l_6$	0

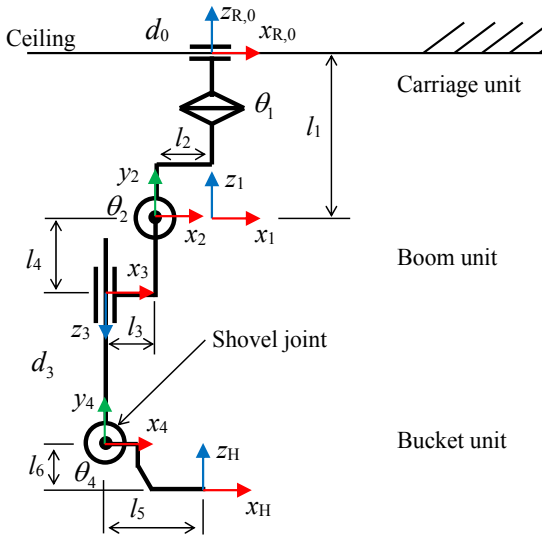

Fig. 5 Mathematical link manipulator model of caisson shovel

3.2 Homogeneous transformation by RGBD sensor information

Figure 6 shows the mathematical model describing the relationship among the working chamber, the caisson shovel, the RGBD sensor, and the terrain. Here, Σ_R, Σ_i, Σ_s, and Σ_C are the coordinate systems for the rail track, the *i*th joint of the caisson shovel, the RGB sensor, and the working chamber ceiling, respectively. Since the RGBD sensor is fixed in the coordinate system, Σ_1, the RGB sensor coordinate system, Σ_s, from the viewpoint of the working chamber ceiling

coordinate system, Σ_C, is calculated by the homogeneous transformation matrix, ${}^{C}T_S$, obtained from the coordinate systems, Σ_C, Σ_R, Σ_0, Σ_1, and Σ_S:

$$ {}^{C}T_S = {}^{C}T_R \, {}^{R}T_0 \, {}^{0}T_1 \, {}^{1}T_S \tag{4} $$

$$ {}^{C}T_R = \begin{bmatrix} \cos\phi_R & -\sin\phi_R & 0 & x_R \\ \sin\phi_R & \cos\phi_R & 0 & y_R \\ 0 & 0 & 1 & z_R \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

$$ {}^{R}T_0 = \begin{bmatrix} 1 & 0 & 0 & d_0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

$$ {}^{0}T_1 = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

$$ {}^{1}T_S = \begin{bmatrix} \cos\phi_S \cos\theta_S & -\sin\phi_S & \cos\phi_S \sin\theta_S & x_S \\ \sin\phi_S \cos\theta_S & \cos\phi_S & \sin\phi_S \sin\theta_S & y_S \\ -\sin\theta_S & 0 & \cos\theta_S & z_S \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

where (x_R, y_R, z_R) and ϕ_R are the origin and yaw orientation of the rail track coordinate system from the viewpoint of the working chamber ceiling coordinate system, Σ_C. Similarly, (x_S, y_S, z_S) and (ϕ_S, θ_S) indicate the origin and the yaw and pitch orientation of the RGBD sensor coordinate system from the viewpoint of first joint coordinate system, Σ_1. Note that the terrain map translated by this equation is saved as digital elevation model (DEM) information [13]. Since the DEM describes the terrain map as a set of heights on the *x*-*y* structured mesh plane (*n*×*m* grid space), the memory size is reduced.

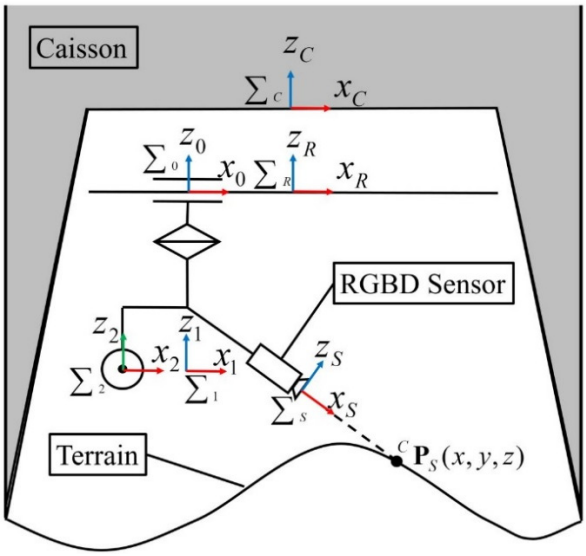

Fig. 6 Mathematical model of working chamber

4 Terrain mapping experiment

In order to estimate the accuracy of this system, we performed a terrain mapping experiment using an artificial ground. **Figure 7** shows the experimental setup consisting of the rail track, the caisson shovel that the boom and bucket units are removed for simplification, and the target terrain. **Figure 8** shows the artificial ground (target terrain) constructed with eight rectangular column blocks having base areas of 100mm×100mm on a horizontally situated flat plate of 1800mm×910mm (Σ_A). **Table 3** lists the X and Y positions and Z heights of the rectangular column blocks. The block configuration described as (X[mm], Y[mm], Z[mm]) in Σ_A is transformed into the DEM data described as (X[grid], Y[grid], Z[mm]).

In this paper, we performed terrain mapping experiments for two cases: Case 1 maps based on information obtained from a fixed caisson shovel, whereas Case 2 maps based on multiple time series information obtained continuously from a caisson shovel traveling on the rail track. Note that the DEM resolution (*n*×*m* grid space) was set to 1300pixels×1000pixels, where the pixel size is 10mm.

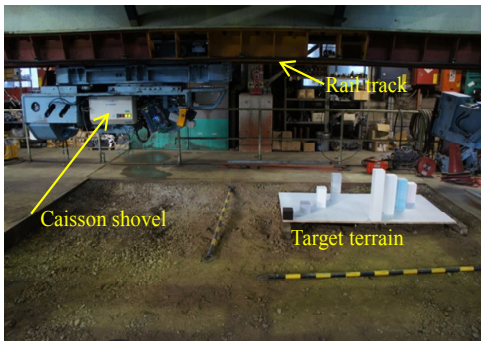

Fig. 7 Experimental setup

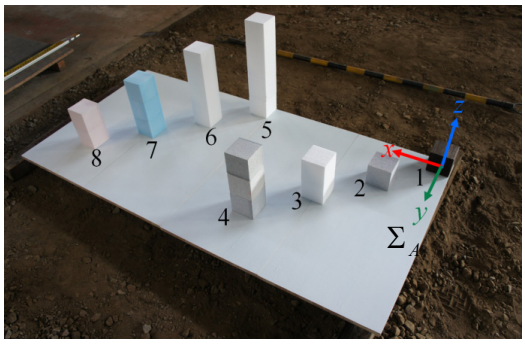

Fig. 8 Artificial terrain with eight rectangular blocks

Table 3 Block position [mm (grid)] and height [mm]

ID	X	Y	Z (height)
1	0 (0)	0 (0)	100
2	200 (20)	200 (20)	100
3	400 (40)	400 (40)	200
4	600 (60)	600 (60)	300
5	900 (90)	0 (0)	500
6	1100 (110)	200 (20)	400
7	1300 (130)	400 (40)	300
8	100 (10)	600 (60)	200

5 Results of terrain mapping and discussion

Figures 9, **10**, and **11** show the RGB, depth, and point cloud images transformed by eq. (4), as an example. In the monochrome depth image, the darker the color, the longer the distance. The outside area from the depth sensor range is depicted as black (= 0.0).

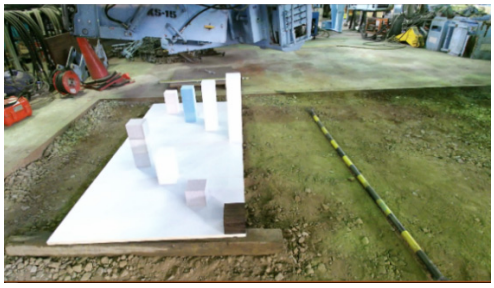

Fig. 9 Example of RGB image

Fig. 10 Example of depth image

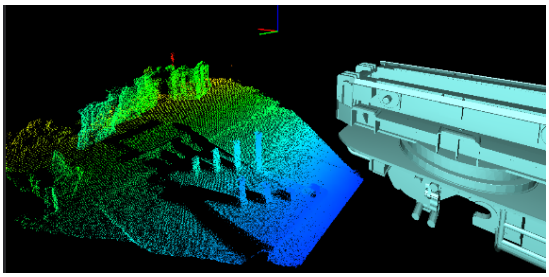

Fig. 11 Example of transformed point cloud image

Figure 12 shows the obtained DEM terrain map for Case 1, in which the caisson shovel, i.e., the RGBD sensor, was fixed. In this figure, high and low altitudes are indicated in red and blue, respectively. The grid area for which the cloud point does not exist is colored gray. Since this map is created from an image, the occluded area is also shown in gray. Note that the DEM data is created thorough a 5×5 median filter and a 3×3 smoothing filter for noise reduction. **Table 4** lists the error compared to the transformed position and height. Here, the X-Y position was estimated from the grid position, i.e., two dimensional array ID and the Z height was calculated as the average heights of the point cloud in one grid area. The average errors of X, Y, and Z were -3.9grid, -1.4grid, -82mm (std. 1.8grid, 1.4grid, 143mm). The large error and deviation result from the lack of cloud points for each grid area by occlusion. Observation information not from one direction but from many directions is needed for accurate mapping. Note that the theoretical accuracy of the depth sensor is 1mm and the practical grid resolution is 10mm/grid. Hence the accuracy between the X-Y position and the Z height differs almost 10 times.

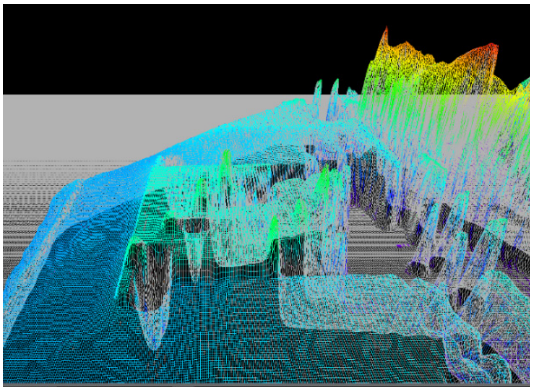

Fig. 12 DEM terrain map obtained for Case 1

Table 4 Calculated block position [grid] and height [mm] error for Case 1

Block ID	X(grid)	Y(grid)	Z (mm)
1	-3	-1	12.1
2	-2	-1	5.5
3	-2	-1	4.0
4	-3	-1	16.6
5	-6	1	-380.8
6	-6	-4	-193.9
7	-3	-2	0.7
8	-6	-2	-123.3

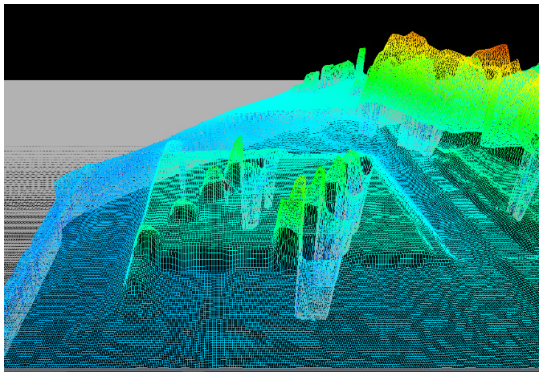

Fig. 13 DEM terrain map obtained for Case 2

Table 5 Calculated block position [grid] and height [mm] error for Case 2

Block ID	X(grid)	Y(grid)	Z (mm)
1	1	-1	1.2
2	1	1	-10.4
3	2	2	-16.7
4	0	2	-36.9
5	-1	-2	-49.9
6	-1	-2	-36.8
7	1	0	-39.6
8	0	-1	-9.3

Figure 13 shows the obtained DEM terrain map for Case 2, in which the caisson shovel, i.e., the RGBD sensor, traveled on the rail track at a velocity of 0.05m/s. This map is integrated from 130 frame images. The DEM data is created thorough the same filters used for Case 1. **Table 5** the error compared to the transformed position and height. The average errors of X, Y, and Z were 0.4grid, -0.1grid, -24mm (std. 1.1grid, 1.6grid, 18mm). Case 2 created a more accurate terrain map than Case 1, because each grid area in the DEM obtained sufficient cloud points. This result suggests that even if the sensor resolution was coarse, multiple time series information and/or information from many sensors improves the accuracy of 3D terrain mapping. Since the errors for Case 2 was just a few centimeters or less and the standard deviation was also small, we conclude that the accuracy of mapping was sufficient for the excavation operation and the reliability was also high. Note that although the blocks in **Figs. 12** and **13** deformed from the rectangular column shape into rounded rectangular column, it is due to normalization through the transformation to the DEM format with a 10mm×10mm resolution. Since the rise and fall of the actual soil is gentle, this deformation can be neglected.

6 Conclusions

In this paper, we developed a 3D terrain mapping method integrating information obtained from a RGBD sensor mounted on the caisson shovel. This system mapped artificial terrain using several columns by translating the cloud points from the camera coordinate system to DEM information from the caisson coordinate system. The accuracy and reliability of the terrain map integrated from multiple information sources was a few centimeters or less and was sufficiently high for the excavation operation.

Mapping integrating information from multiple depth sensors, soil property estimation by integration of RGB images, and motion planning for automatic excavation are our future works.

Acknowledgements

The present research was supported by a grant through the “Sentan Sangyo Souzou Project” of Saitama Prefecture and by Sankeikogyo Corp. The authors would like to express their deep gratitude to all involved in the research project.

References

[1] http://www.orsc.co.jp/english/tec/newm_v2/ncon02.html.

[2] Kodaki, K., Nakano, M., and Maeda, S., “Development of the automatic system for pneumatic caisson”, Automation in Construction, Vol. 6, (1997), pp. 241-255.

[3] Eisenbeiß, H., “UAV Photogrammetry, Dissertation Institute of Geodesy and Photogrammetry”, ETH Zurich, Switzerland, 2009.

[4] Remondino, F., Barazzetti, L., Nex, F., Scaioni,M., and Sarazzi, D., “UAV photogrammetry for mapping and 3Dmodeling–current status and future perspectives”, Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g) 2011, Zurich, Switzerland, September 2011.

[5] Zongjian, L.I.N., “UAV for mapping—low altitude photogrammetric survey”, International Archives of Photogrammetry and Remote Sensing, Beijing, China, 2008.

[6] Hudzietz, B.P., and Saripalli, S., “An experimental evaluation of 3d terrain mapping with an autonomous helicopter”, Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g) 2011, Zurich, Switzerland, September 2011.

[7] Bulatov, D., Solbrig, P., Gross, H., Wernerus, P., Repasi, E., and Heipke, C., “Context-based urban terrain reconstruction from UAV-videos for geoinformation applications”, Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g) 2011, Zurich, Switzerland, September 2011.

[8] Neitzel, F., and Klonowski, J., “Mobile 3D mapping with a low-cost UAV system”, Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. 38 (2011) 1–6.

[9] Su, Y.Y., Hashash, Y.M.A., and Liu, L.Y., “Integration of Construction As-Built Data Via Laser Scanning with Geotechnical Monitoring of Urban Excavation”, Journal of Construction Engineering and Management, Volume 132, 2006.

[10] Duffell, C.G., and Rudrum, D.M.,_2005. “Remote sensing techniques for highway earthworks assessment.” Proc., Geo-Frontiers 2005.

[11] Yamamoto, H., Uesaka, K., Ishimatsu, Y., Yamaguchi, T., Aritomi, K., and Tanaka, Y., “Introduction to the General Technology Development Project: Research and Development of Advanced Execution Technology by Remote Control Robot and Information Technology”, ISARC, 2006.

[12] Yamamoto, H., Ishimatsu, Y., Ageishi, S., Ikeda, N., Endo, K., Masuda, M., Uchida, M., and Yamaguchi, H., “Example of Experimental Use of 3D Measurement System for Construction Robot Based on Component Design Concept”, ISARC, 2006.

[13] Yoshimitsu, T., Kubota, T., and Nakatani, I., “Path Planning for Exploration Rovers over Natural Terrain Described by Digital Elevation Map”, Journal of the Robotics Society of Japan, Vol. 18 No. 7, 2000.

Received on December 29, 2016
Accepted on March 29, 2017