Article

3D Pose Recognition System of Dump Truck for Autonomous Excavator

Ju-hwan Lee [1], Junesuk Lee [2] and Soon-Yong Park [3,*]

1   School of Computer Science and Engineering, Kyungpook National University, 80 Daehak-ro, Puk-gu,
     Daegu 41566, Korea; sasinhwan@naver.com
2   School of Electronics and Electrical Engineering, Kyungpook National University, 80 Daehak-ro, Puk-gu,
     Daegu 41566, Korea; ljs2352@gmail.com
3   School of Electronic Engineering, Kyungpook National University, 80 Daehak-ro, Puk-gu, Daegu 41566, Korea
*   Correspondence: sypark@knu.ac.kr; Tel.: +82-053-950-7575

Abstract: The purpose of an excavator is to dig up materials and load them onto heavy-duty dump trucks. Typically, an excavator is positioned at the rear of the dump truck when loading. In order to automate this process, this paper proposes a system that employs a combined stereo camera and two LiDAR sensors to determine the three-dimensional (3D) position of the truck's cargo box and to analyze its loading space. Sparse depth information acquired from the two LiDAR sensors is used to detect points on the door of the cargo box and establish the plane on its rear side. Dense depth information of the cargo box acquired from the stereo camera sensor is projected onto the plane of the box's rear to estimate its initial 3D position. In the next step, the point cloud sampled along the long shaft of the edge of the cargo box is used as the input of the Iterative Closest Point algorithm to calculate a more accurate cargo box position. The data collected from the stereo camera are then used to determine the 3D position of the cargo box and provide an estimate of the volume of the load along with the 3D position of the loading space to the excavator. In order to demonstrate the efficiency of the proposed method, a mock-up of a heavy-duty truck cargo box was created, and the volume of the load in the cargo box was analyzed.

Keywords: autonomous excavator; 3D pose recognition; dump truck

1. Introduction

There is a shortage of new construction workers in the industry due to the trend of laborers wanting to avoid extreme environments [1]. Particularly, due to the retirement of older workers, there is a shortage of skilled workers. Therefore, research is being conducted on autonomous trucks and construction robots that can replace humans at construction sites. Excavators are considered low risk because they do not travel at high speeds, and they perform repetitive tasks. Hence, their automation is highly feasible. For this reason, major construction equipment manufacturers are incorporating various types of automation [2,3]. Diverse research was carried out in line with this trend. For example, several publications examined the automated determination of excavation spaces and related control techniques [4–11]. An excavator digs up and transports loads into the cargo box of heavy-duty dump trucks. Excavators feature long joints and are often positioned at the rear of a truck to load materials. After the cargo box has been loaded sufficiently, the process is repeated for additional trucks.

To automate this process, it is vital that the excavator acquire accurate information about the cargo box's three-dimensional (3D) position for analysis. Stentz introduced a method for recognizing a truck's cargo box after segmenting a truck from dense Lidar data [12]. The truck's cargo box is scanned using a dense LIDAR sensor. After segmenting the cargo box from dense rider data, the 3D position of the dump truck is determined by the upper plane of the cargo box. However, suppose the excavator is positioned higher than the

truck, and the truck is parked to the excavator's side. Stereo cameras are efficient sensors that provide dense 3D depth maps. J. R. Borthwick introduced technology for recognizing the location of a truck using depth information obtained from a stereo camera [13]. The 3D pose is estimated using the 3D shape information of the truck. A truck cargo box is formed from several faces. The 3D position of the truck is calculated by matching all cargo box planes and the plane of the initial 3D model. Therefore, 3D shape information of the entire truck is required. This method can be used for very large excavators, but it is difficult to apply to general excavators with a low height.

Moreover, there is a way to load a load into a truck by using the structure after installing a special structure [14]. Since the truck is located inside the special structure and the excavator knows the 3D position of the special structure, the truck's location is naturally estimated. This method allows the excavator to locate the truck's 3D location reliably, but it is costly and limited in the location of the structure.

Another option is to place the sensor in a third location other than excavators and trucks [15]. After installing the GNSS sensor that indicates the location of the truck and excavator, the initial 3D location of the truck and excavator is expressed in a global coordinate system. Next, the three-dimensional position of the truck is estimated using four lidar sensors and fish-eye cameras installed on the excavator. This method increases the cost of configuring the sensor.

Despite the wide variety of dump truck types available, their cargo boxes generally resemble hexahedrons. Furthermore, the box's door is a flat plane shape for simplified fabrication. This paper describes a technique that enables an excavator at the rear of the loading truck to recognize its 3D cargo box and to estimate the 3D position of the loading space and the volume of the load inside using data obtained from a novel sensing device assembled for this purpose.

Specifically, this research provides the following three innovations:

1. A novel dump truck cargo box sensing device;
2. A method of automatically estimating the 3D position of the loading space;
3. A method of automatically estimating the volume of the load.

The sensing device combines two two-dimensional (2D) LiDAR sensors and a stereo camera. The scenario includes an excavator operating in an outdoor environment. Note that using only a stereo camera would be hazardous. Hence, two LiDAR sensors complement it (Section 2). The distance values acquired by each sensor are projected onto one coordinate system using a calibration board and a line laser for accuracy (Section 3). As the cargo box door resembles a plane shape, the 3D estimation of the position of the cargo box is achieved by determining the plane of the rear of the cargo box from the data received from the two LiDAR sensors installed vertically on the sensing device (Section 4.1). Furthermore, the 3D position of the box is determined by projecting the dense 3D location of the door acquired from the camera onto a plane and matching it with the actual 3D model of the cargo box. The 3D position of the box can be obtained by utilizing one virtual matching point and four matching points of the rear plane (Section 4.2).

If the 3D position of the cargo box is estimated solely using the plane, errors may occur. Therefore, points along the long shaft of the box are used for more accurate determination. By matching the sampling of the points of the box's long shaft to the actual 3D model, the distance data of the shaft obtained from the stereo camera are fed to an iterative closest point (ICP) algorithm (Section 4.3). Lastly, after the 3D position of the cargo box is estimated, the volume of the load is calculated. Moreover, the 3D position information for the loading space is transmitted to the excavator.

The general outline of this paper is as follows. First, Section 2 describes the sensing device developed for this study. In Section 3, the method of calibrating the sensors of the sensing device is discussed. Section 4 outlines the method for estimating the truck cargo box using information acquired from the sensing device. Section 5 examines the accuracy of the proposed method for estimating load volume and determining the loading space. The study concludes with Section 6.

2. Sensor Configuration

Figure 1 illustrates the sensing device used to detect the dump truck cargo box. Its dimensions were 20 × 10 × 30 cm in terms of width, length, and height, respectively. The device consisted of a stereo camera (ZED-1) and two LiDAR (Hokuyo UST-10LX1-CH) sensors and was placed on top of the excavator's driver seat to estimate the plane of the cargo box's rear. The stereo camera provides one color image and one 3D depth map. The data acquired by the stereo camera were used to recognize the 3D position of the dump truck's cargo box, determine the loading space available, and calculate the volume of the load. Because most excavators operate outdoors, a commercial outdoor camera was selected. It had a resolution of 1920 × 1080 pixels and provided red green blue video information as a depth map at a rate of 30 fps. The depth map had sufficient sensitivity between 0.5 and 20 m.
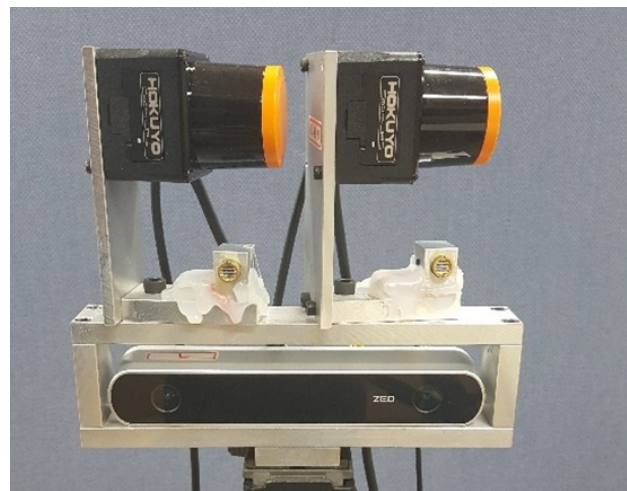

Figure 1. Dump truck cargo box 3D position recognition system.

An excavator's working environment comprises dust and solar reflections. Consequently, the information acquired from a camera in such an environment contains noise; hence, errors occur in the distance estimates of the stereo algorithm. Therefore, for the purposes of this research, it was not possible to rely solely on the information captured by the camera. The two LiDAR sensors were attached perpendicular to the device. They are widely used with automated driving technologies, drones, and robots. One disadvantage of LiDAR is its price, which increases with resolution. Although LiDAR acquires sparse depth information, it has the advantage of consistent accuracy. When combined with camera data, superior dense depth information is acquired. The LiDAR sensor has a 270° angle and can see distances between 0.02 and 10 m. It provides users with a relatively high level of depth accuracy at ±40 mm (Table 1).

Table 1. Specifications of the proposed dump truck 3D position recognition system.

Function	Component	Manufacture	Model	Description
3D Sensing Device	LIDAR	HOKUYO	UST-10LX1-CH	Field of view: 270°
				Depth Range: 0.02–10 m
				Accuracy: ±40 mm.
	Camera	STEREOLABS	ZED 1	Resolution: 1920 × 1080
				Depth Range: 0.5–20 m

3. Sensor Calibration

During the scanning process, the depth information from each sensor can be retrieved in real-time. However, the 3D depth information depends on the coordinate system of the sensor. Therefore, it is essential to define a method for representing depth information in a single coordinate system via sensor calibration. For ease of calibration, the line lasers and a flat calibration object were used, as shown in Figure 2.

(a)

(b)

Figure 2. Calibration tool to calibrate the sensor. (**a**) calibration object; (**b**) line laser.

Generally, low-cost, mass-produced LiDAR sensors have a bandwidth of 905 nm. The scanning area with the infrared (IR) filter removed captures the laser lines [16]. Hence, when the line laser and LiDAR scan areas are matched, the area to be scanned by LiDAR can also be determined from the stereo camera image, as shown in Figure 3.
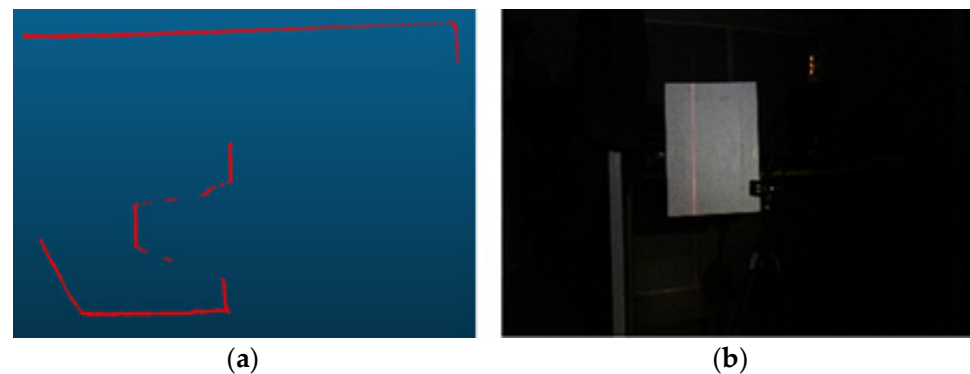

(a)

(b)

Figure 3. Example of data acquisition process using a calibration plate. (**a**) Depth information obtained from LiDAR; (**b**) Image information obtained from stereo camera.

With T defined as the Euclidean transformation relation between two coordinate systems, the transformation matrix, T, is represented by a single (3×3) rotation transformation matrix, R, and a (3×1) translation transformation vector, t. When the 3D matching points obtained from the stereo camera and the LiDAR sensor are expressed as Q_i and P_i, respectively, and three matching points theoretically exist, the transformation matrix, $T_{(L \rightarrow C)}$, between the two coordinate systems can be obtained as follows:

$$Q_n = T_{L \rightarrow C} P_n \tag{1}$$

If there are more than three matching points, the transformation matrix between the two points in time can be obtained using the singular value decomposition (SVD) or the Levenberg–Marquardt (LM) algorithm to minimize the error [17], as shown in Equation (2):

$$\varepsilon = \sum_n || Q_n - T_{L \rightarrow C} P_n ||. \tag{2}$$

In order to facilitate calibration, the LiDAR sensors were positioned so that they sensed the calibration board, and the center point of the line was detected by each LiDAR

sensor and camera as a match point, as shown in Figure 4. The distance information of the calibration board scanned by the LiDAR sensor was segmented as a distance value for detection, and the stereo camera used the images to detect the calibration board. Based on the random sample consensus algorithm (RANSAC), the detected lines were fitted more precisely, and the center point was determined [18].
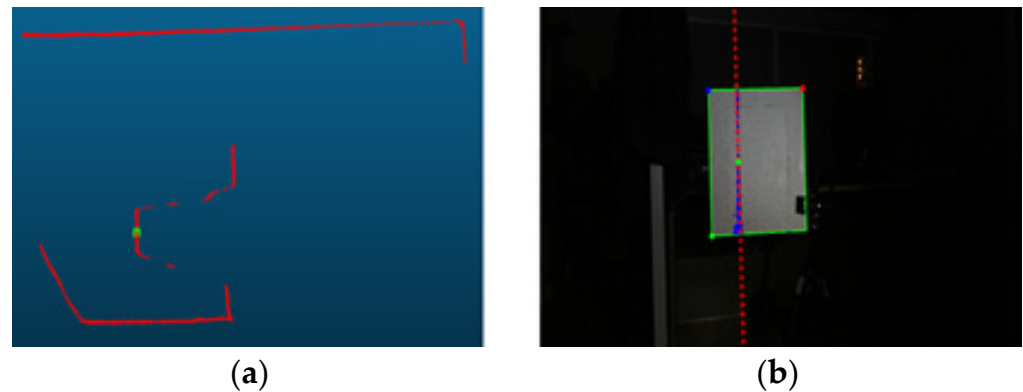

Figure 4. Example of detecting the center point from the depth information acquired by LIDAR and stereo. (**a**) A center point of the line in LIDAR data; (**b**) A center point of the line in the camera data.

After placing the calibration board at various positions and obtaining the transformation matrices using Equation (2), an accurate projection of the data obtained from the LiDAR sensor onto the position of the line laser was available, as shown in Figure 5. The accuracy of calibration becomes greater with the increasing number of match points, but this study confirmed that six were sufficient. Figure 6 shows the results when inter-device calibration was performed and when it was not.
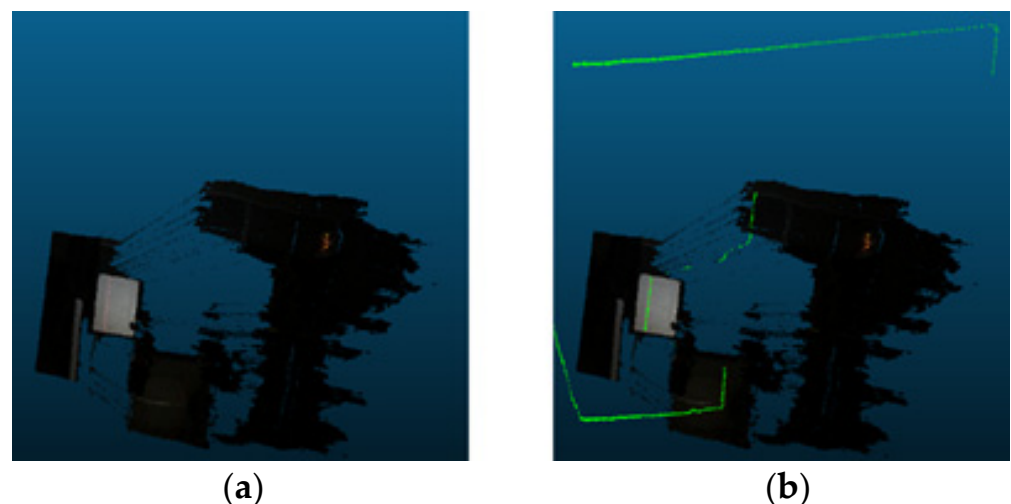

Figure 5. The result of calibrating the LIDAR and stereo camera using the center points. (**a**) 3D point cloud obtained from the stereo camera; (**b**) The result of superimposing the information scanned from the LIDAR on the stereo 3D point cloud.

With the calibrated device in the outdoors scan experiment, it is possible to obtain three-dimensional depth information expressed in one coordinate system. In this study, the reference coordinate system of the device was defined as the stereo camera's coordinate system.
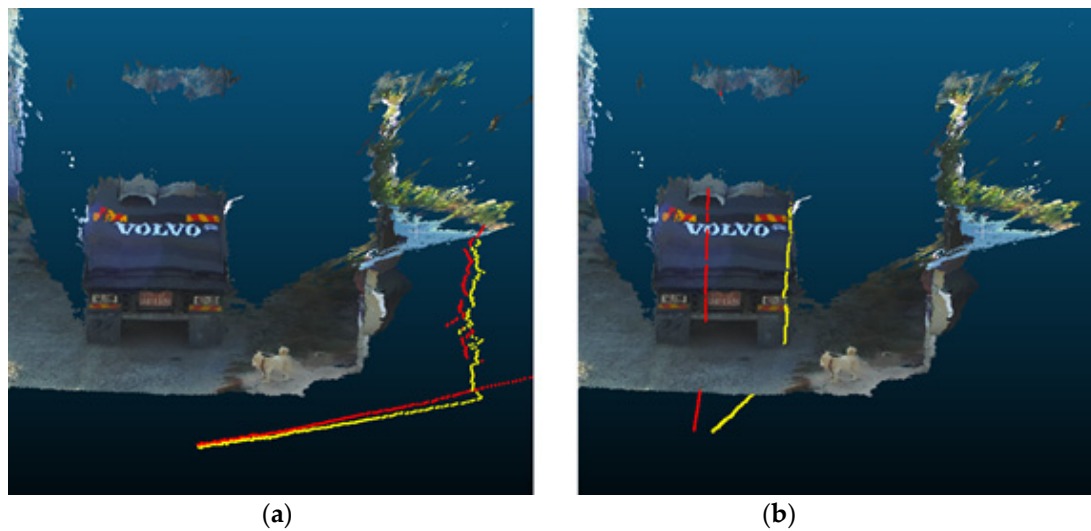
| (**a**) | (**b**) |

Figure 6. The result of projecting data from two LiDAR sensors on a dense 3D point cloud obtained from a stereo camera. The red line is the left LIDAR sensor, and the yellow line is the right LIDAR sensor. The left side of the sensor: (**a**) the result before correction; (**b**) the result after correction.

4. Location Estimation of Truck Cargo Box

In order to load the excavator's content into the truck's cargo box, it is first necessary to determine its 3D position. Heavy-duty dump truck cargo box doors are flat and generally mounted on the rear of the box. It is characterized by a long shaft emerging from its edge. In this research, the accuracy of the location of the plane was prioritized, assuming the device senses the cargo box door of the truck and that the rear of the cargo box is flat.

In Figure 7, a flowchart of the method of estimating the 3D position of the truck cargo box is presented. After the rear of the cargo box is scanned by the device, a dense and sparse 3D point cloud is generated by the stereo camera and the two vertically mounted LIDARs, respectively. Based on the feature points associated with the images on the left and right, the stereo algorithm generates the 3D depth value. Thus, the more distinct feature points detected by the left and right cameras, the more accurate the 3D depth [19]. Meanwhile, the depth value of a single-color object with few features will have a larger error because of inaccurate matching. It is possible to scatter random patterns to overcome this problem, but they are ineffective in the presence of strong wavelengths of light, such as sunlight [20,21].
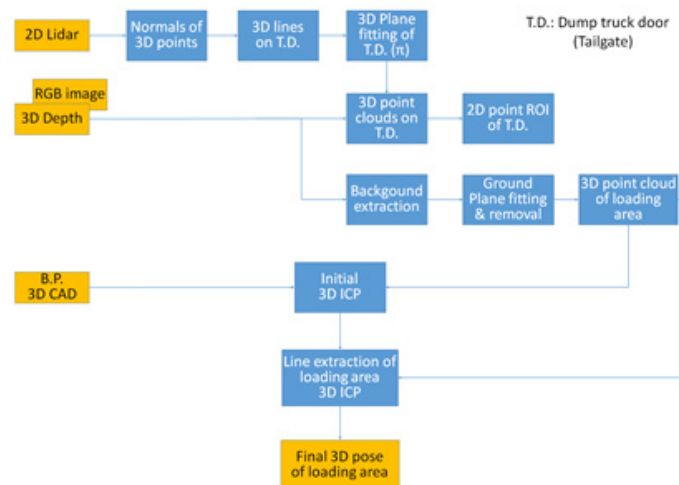

Figure 7. Flow chart of 3D pose localization processing of truck cargo box.

Commercial dump trucks usually have a single-color cargo box, making it difficult for left and right cameras to identify similar features. Consequently, the distance information on the rear side of the dump truck cargo box will be inaccurate. To overcome this issue, two LiDAR sensors that provide constant depth information were used to identify points on the cargo box's door. The normal vector of the 2D point first obtained from the lidar sensor is calculated using the surrounding points. The points on the door are determined by the distance information between the normal vector and the two LiDAR sensors, and the noise points are removed using a line fitting algorithm. They are then formulated into a plane corresponding to the rear of the cargo box.

A stereo camera provides a dense 3D point cloud. Points close to the plane are points on the door. These points are projected onto a plane and represented as a 2D ROI. Conversely, points that do not exist on the plane are considered noise. These 3D points exist on the ground and are removed using ground plane fitting. As a result, a 3D point cloud of the truck cargo box is obtained.

In the next step, we create one 3D hexahedron model with the same dimensions as the cargo box. The initial 3D position of the cargo box is determined by matching the planes of the two models. After estimating the initial 3D position of the box, the estimation error is corrected by using the long axis of the cargo box. A detailed description of each step can be found in the detail section.

4.1. Cargo Box Rear Detection Using a LiDAR Sensor

Due to the vertical placement of the two LiDAR sensors, data from the cargo box door and the ground are merged, and the point cloud for the cargo door is separated. In theory, a 3D point cloud on a plane can be described by an identical normal vector. Furthermore, when a sensing device is constructed, the index, i, of the point cloud obtained from the two LiDAR sensors may be considered identical if both identical LiDAR sensors are located on the same plane. Therefore, a simple method was used to detect the point cloud on the cargo box door using two LiDAR sensors.

Figure 8 illustrates the process of obtaining the point cloud on the cargo box door. Three-dimensional point clouds acquired from the left and right LiDAR sensors are denoted as P_i^L and P_i^R And their normal vectors are $\hat{P}_i^L$ and $\hat{P}_i^R$, respectively. The normal vector, $\hat{P}_i^L$ is determined by the cross product of the two vectors after obtaining the direction vector, B P_i^R of P_i^L and P_{i+k}^L (k is 3–5), after the direction vector, A of P_i^R, is ascertained from, P_i^L. The normal vector, $\hat{P}_i^R$, of the right-side LiDAR data, P_i^R, is obtained using the same method. When the normal vector of all point clouds is determined, the dot product and the z-axis (0, 0, 1) of the stereo camera coordinate system are obtained.
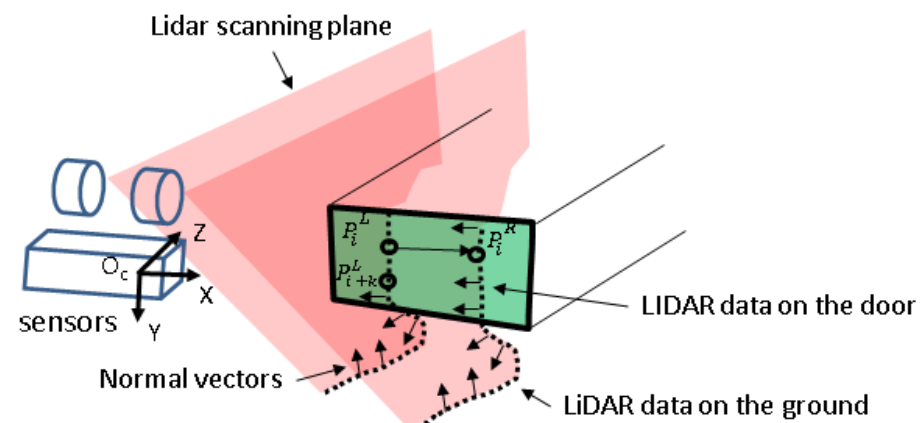

Figure 8. A method to detect the door of a dump truck cargo box using two LIDAR sensors. The points on the plane have the same normal vector, and the dot product with the camera z-axis is close to −1. The normal vector is computed with two LIDAR point clouds.

As the sensor mounted on the excavator is placed on the rear side of the cargo box, an inner product value closer to -1 is considered to be the point cloud on the cargo box door. The point cloud is filtered using the magnitude of the direction vector, A. Points with significant differences in the distance cannot be considered part of the point cloud and are therefore removed, as they are considered noise. Because the points are discriminated using only the normal vector and the distance between the point clouds, there may be outliers that are not removed from the filtering.

Lastly, outliers are eliminated by line modeling the 3D point cloud sorted using the random sample consensus algorithm, and the final 3D point cloud existing on the cargo box door is sorted. Figure 9 shows a LiDAR point cloud with outliers removed from a three-dimensional point cloud of a stereo camera. Point clouds that are not line-fitted are considered noise.
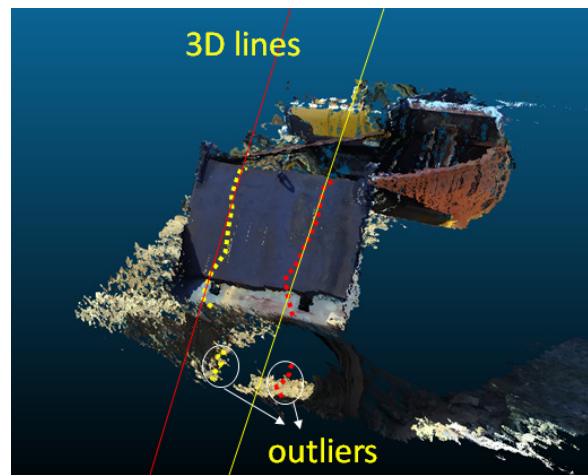

Figure 9. Result of removing outlier point clouds in LiDAR data.

This point cloud is used to define the plane of the rear of the cargo box, π_L, as in Equation (3), and the plane equation is obtained using the least-square method shown in Equation (4):

$$ax + by + cz + 1 = 0. \tag{3}$$

$$\begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ \vdots & \vdots & \vdots & \\ x_i & y_i & z_i & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4}$$

The matrix form is a linear equation, $AX = 0$. Therefore, the general solution, X, is computed using SVD, as in Equation (5). Finally, the back plate plane is computed with Equation (6):

$$A = U \sum V^{T \cdot}, \tag{5}$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = V^{T \cdot} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \pi_L = \begin{bmatrix} a & b & c & 1 \end{bmatrix}^T. \tag{6}$$

4.2. Recognition of Initial 3D Position Using a Stereo Camera

Although the box is plane-shaped, the 3D distance provided by the stereo camera does not provide an accurate representation of the shape if the box is single-colored. Consequently, the plane data obtained by the LiDAR sensor on the rear of the cargo box are used to correct the acquired depth value. In the dense depth map obtained from the stereo camera, the 3D point cloud is expressed as $Q_i(x_i, y_i, z_i)$. Hence, the points closer

to the cargo box plane must be sorted in advance as they are likely to be situated on the cargo box door. It is possible to calculate the distance between point Q_i and plane π_L using Equation (7):

$$d = \frac{|ax_i + by_i + cz_i + \mathrm{d}|}{\sqrt{a^2 + b^2 + c^2}}, \, d < 30 \tag{7}$$

Generally, only points with a difference in distance within 30 cm are considered cargo box door points; all other points are regarded as noise and are eliminated. The sorted points can also be projected onto plane π_L using Equation (8):

$$Q'_i = Q_i - d \times \pi_L . \tag{8}$$

Figure 10 shows a projection of the dense 3D point cloud acquired by the stereo camera on the cargo box plane obtained by LiDAR sensors. The distance errors are calibrated as the LiDAR sensor measurement values are projected onto the plane. The projected 3D points can be used to estimate the initial 3D position of the excavator cargo box. The 3D hexahedron model based on the actual dimensions of the box is generated, as shown in Figure 11. Next, four points near the vertex in the projection points cloud are detected to match the hexahedron model to the actual cargo box plane, as shown in Figure 12. The points corresponding to the hexahedron model are determined in the same order. Figure 13 illustrates the method of matching the 3D hexahedron model and the truck cargo box.
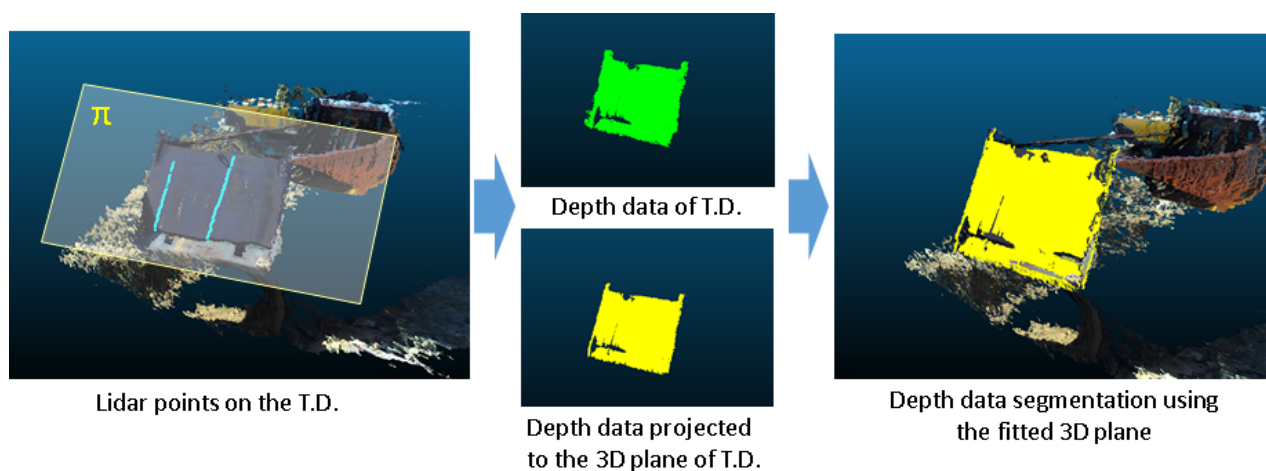

Figure 10. The result of projecting a dense point cloud obtained with a stereo camera onto the equation of the truck's back plate plane obtained from LiDAR.
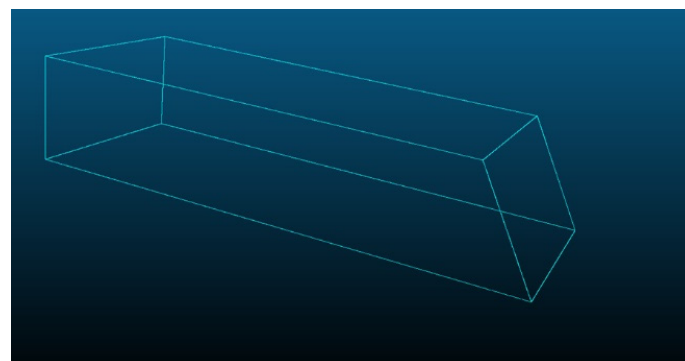

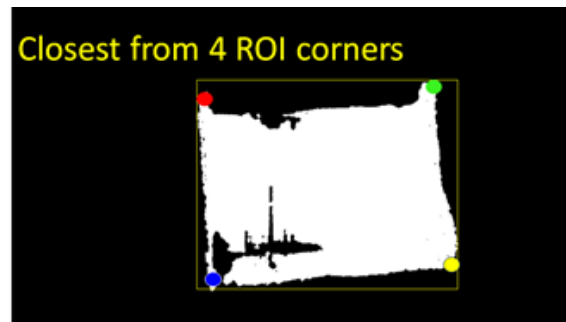**Figure 11.** Three-dimensional model of the cargo area.

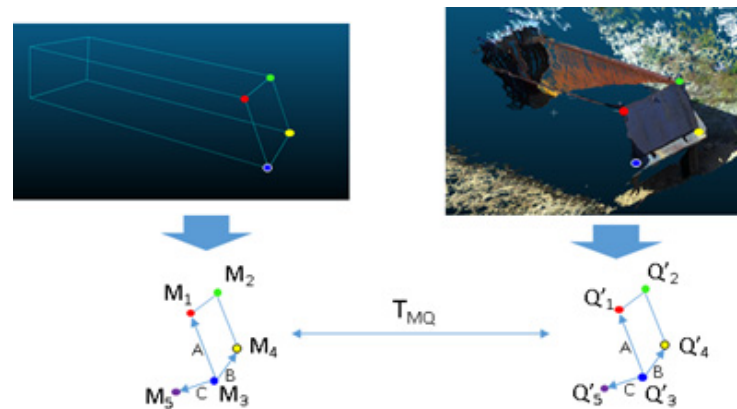**Figure 12.** Example of detecting four vertices of a door.

Figure 13. A method to match the plane between two models.

First, the four vertices (M_1, M_2, M_3, M_4) on the 3D hexahedron model and the four vertices (Q'_1, Q'_2, Q'_3, Q'_4) on the cargo box plane are set as points of convergence. (Q'_5, M_5) is a virtual point of convergence used to eliminate the symmetry ambiguity. The cargo box has a rectangular parallelepiped form, and its calculation is subject to errors. The virtual point of convergence is determined from the cross-product of each point cloud unit vector, $|A|$ and $|\hat{B}|$, which become direction vectors of identical magnitude. When the five points of convergence are defined, the 3D transformation matrix, $T_{M \to C}$, is obtained with its error minimized, as in the following equation. Figure 14 shows the result of matching the cargo box to the hexahedron model using the transformation matrix. The figure shows errors, but a match was made based on the cargo box plane.
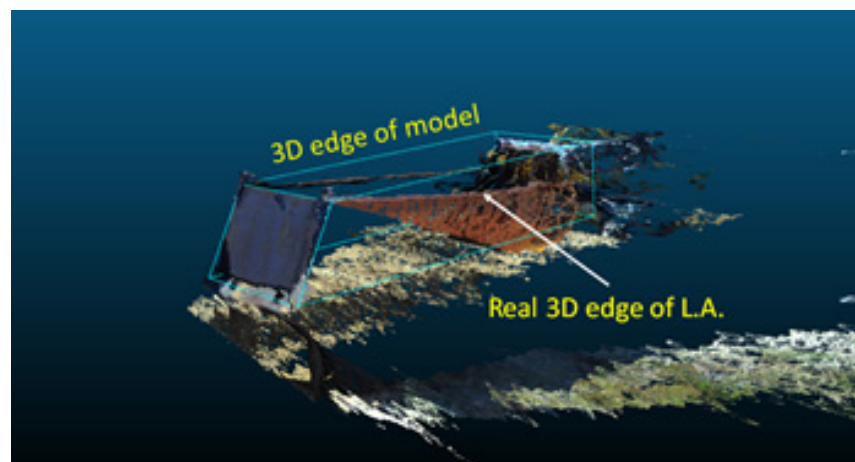

Figure 14. Example of initial 3D position and error of the truck cargo box.

4.3. Pose Estimation Refinement

The truck cargo box position was estimated with only five points of convergence results in the errors presented in Figure 14. This is caused by differences in the distance information estimated by the stereo camera and the generated 3D hexahedron model. The initial matching results show that more errors occurred at the longitudinal corner than at the rear of the cargo bed. Therefore, to minimize error, the points of convergence are sampled at regular intervals from the two corners along the longitudinal axis of the 3D hexahedron model, as shown in Figure 15. An iterative algorithm is used so that these points of convergence and the longitudinal axis of the cargo box obtained by the camera will match.
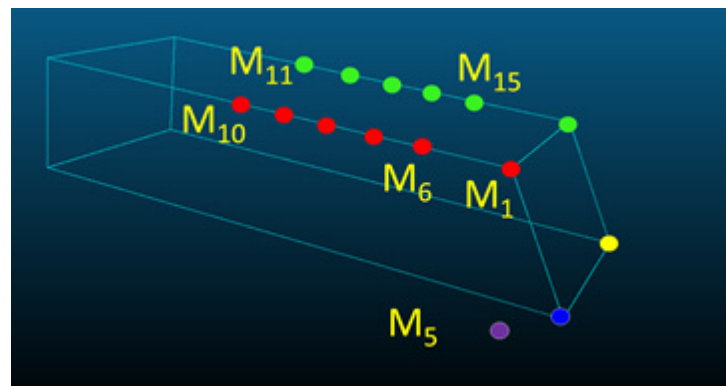

Figure 15. Example of 5 3D points sampled for each corner along the long axis of a cargo box.

A 3D transformation matrix, $T_{M \to C}$, is used to project the sampled 3D point, M_i, of the corner onto the camera's 2D image using Equation (9):

$$m_i = k\, T_{M \to C} M_i. \tag{9}$$

Here, k is the internal parameter of the stereo camera, and m_i is the point projected onto the 2D camera image, as shown in Figure 16. If the initial position transformation matrix, $T_{M \to C}$, is accurate, the projected points are identical to the box corner. Otherwise, they do not match, as shown in Figure 16. Thus, the longitudinal axis of the cargo box is used to solve this issue. Figure 17 presents the method for detecting points of convergence. When they are projected onto the 2D image, the left corner of the cargo box is searched from the right side of the x-axis, and the right corner is searched from the left side. When a 3D value exists on the searched point, it is designated as the point of convergence. The initial transformation matrix, $T_{M \to C}$, is refined after the ICP algorithm iteratively minimizes the errors of the points added to the corner and the plane [22,23]. The pseudocode of the iterative refinement is shown below:

$T_{M \to C} = InitializeMatrix$
  $While(\varepsilon > T_\varepsilon)$
// Find the corresponding points m'_i where the distance between
    m_i and m'_i is minimum.
    $m'_i = ProjectionPoint(T'_{M \to C}, M_i)$
// Find the T matrix that minimizes the error.
    $\varepsilon = \sum_n || m'_i - k\, T'_{M \to C} M_i ||$
    $T'_{M \to C} = T_{M \to C} T'_{M \to C}$
// convert M_i to global coordinate system.
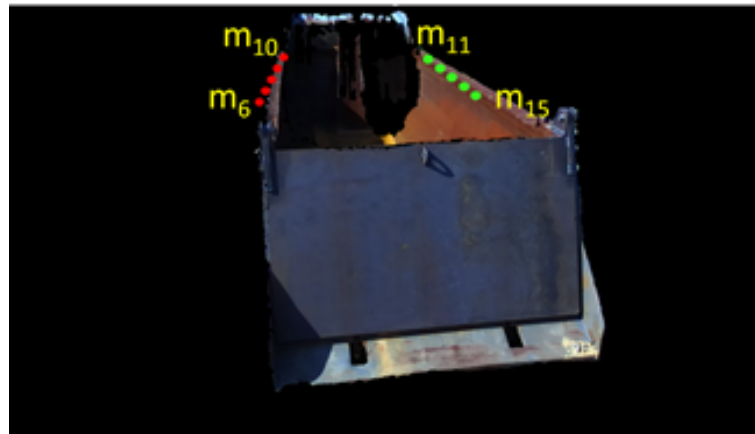 $M_i = T'_{M \to C} M_i$

Figure 16. Example of the 3D points of a 3D model projected onto a 2D image of the camera.
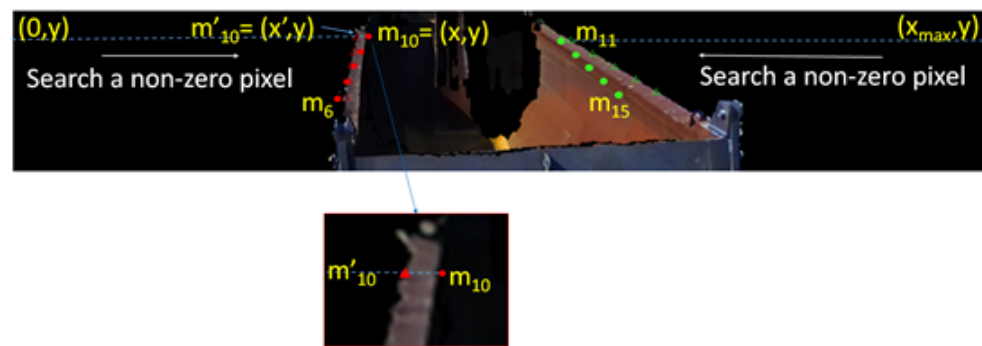

Figure 17. A method of selecting the correspondence point of the truck cargo box from the corners.

5. Results and Analysis

In order to validate the performance of the proposed system, a model of the cargo box of a common heavy-duty dump truck was developed, followed by outdoor experiments. The excavator sensor was installed on the upper section of the driver's seat of the excavator as shown in Figure 18, and the loading procedure was repeated 20 times as the sensor sensed the cargo box while the excavator was moving. The data, as shown in Figure 19, were obtained each time, and the distance values were derived from the stereo sensor.
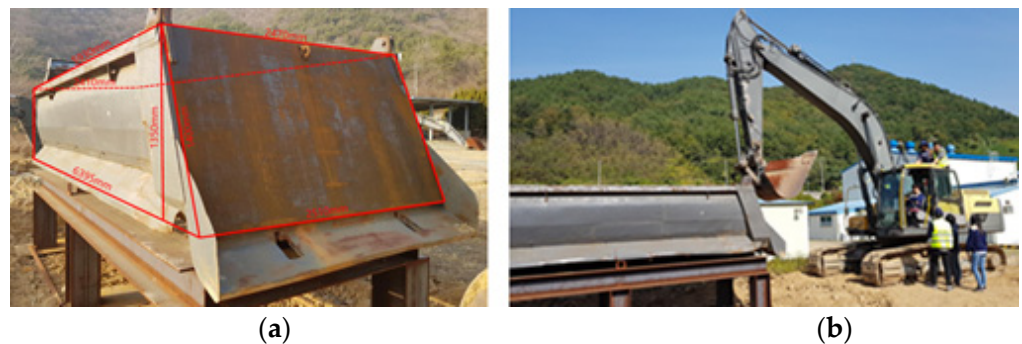

(**a**)   (**b**)

Figure 18. Three-dimensional location estimation experiment for the cargo box of dump truck: (**a**) an experimental model built for experimentation; (**b**) an experiment example using an excavator. The sensor is placed on the top of the excavator.
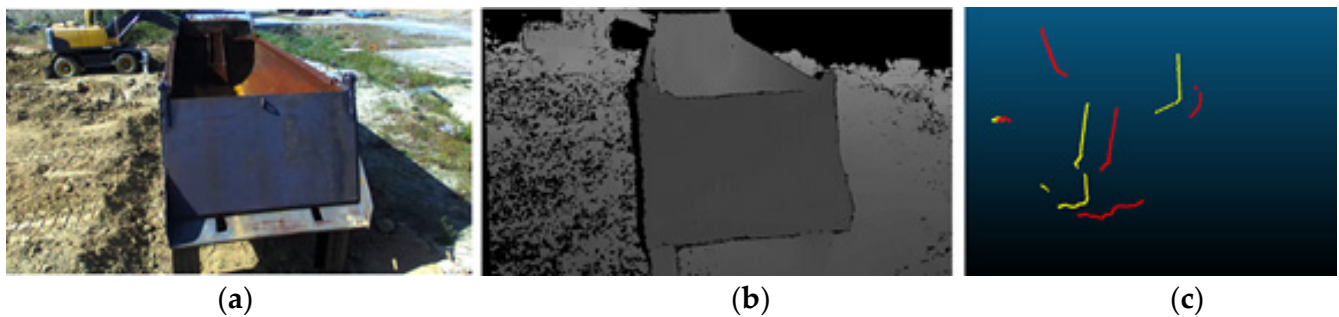
Figure 19. Sensing data obtained from the sensor: (**a**) RGB Image; (**b**) Depth Map; (**c**) LIDAR data.

Figure 20 illustrates the detectability of the truck cargo box door based on distance measurements from the LiDAR sensor. When the distance data obtained by the LiDAR sensor and the stereo camera were visualized, only those points that were estimated to represent the cargo box door were visible. Notably, the widely used dump truck cargo boxes of Volvo and Scania were sensed, excluding the experiment model, and an experiment for detecting the cargo box door was conducted [24,25]. Using the plane data obtained from the LiDAR sensor, the 3D point cloud obtained from the camera was projected. The cargo box doors of two regular trucks are also detected

Figure 20. The result of detecting the doors of two commercial dump trucks with LiDAR data. The green point cloud is the point cloud obtained from the stereo camera projected onto the dump truck door plane.

Figure 21 shows the experimental results of 3D location estimation using the experimental model. The initial results are indicated in blue, and the more precise results using a refined transformation matrix are expressed in yellow. When the initial position was estimated only using plane data, numerous differences were found in the longitudinal axis of the cargo box. The 3D position estimation using the values obtained by the refined transformation matrix showed that the longitudinal axes of the cargo box and the hexahedral model were similar, indicating that the 3D location estimation was successful. Estimation of cargo space volume can increase work efficiency [26,27]. The 3D location of the estimated cargo box enabled the provision of data of the loading space of the cargo box. Figure 22 illustrates the visualized results of normalized data based on box height. The load can be placed in the cargo box, excluding the part in red.
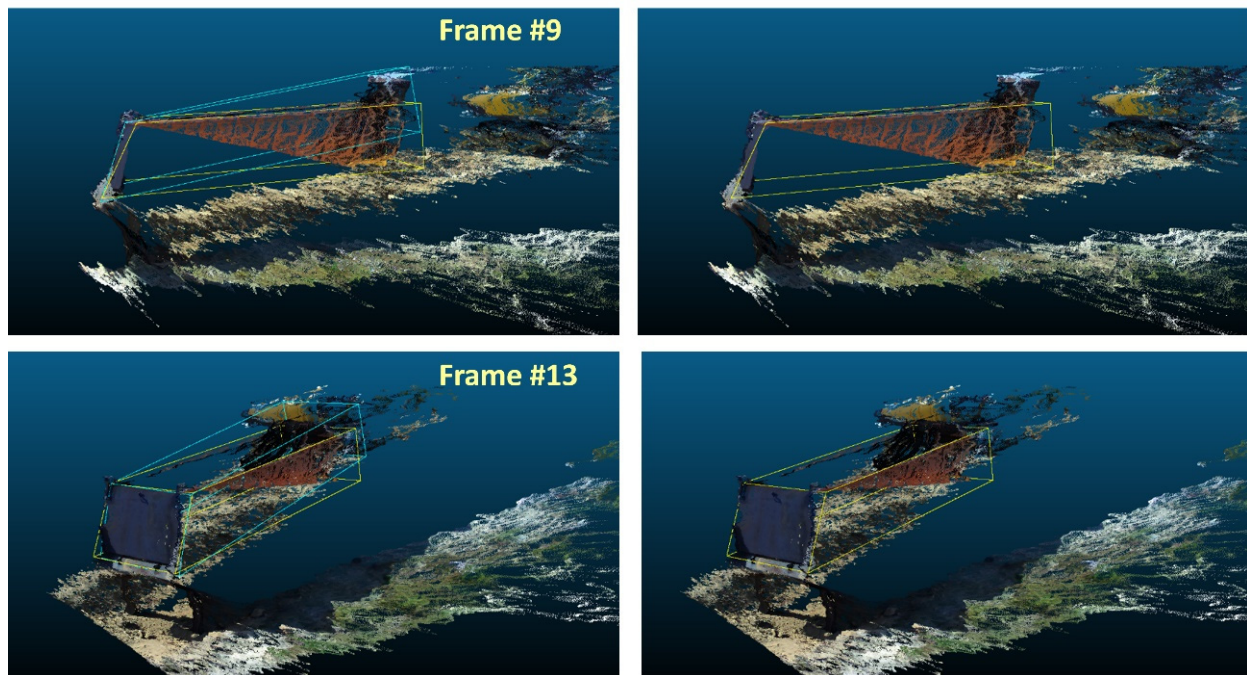
Figure 21. The result of precisely finding the three-dimensional position of the truck cargo box using the proposed method. Blue is the initial 3D position of the cargo box, and yellow is the result of a more precise 3D position. The side corners of the cargo box are more precisely matched.
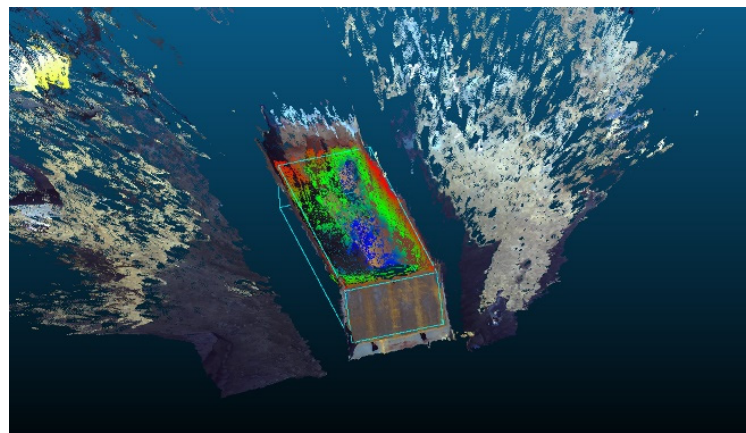

Figure 22. Visualization of the inside of the loader analysis. The excavator can load loads in areas other than the red area.

In this study, the load was loaded five times to the cargo box interior using the excavator, as shown in Figure 23, and volume estimations were made using the image data obtained from the camera. The cargo box interior was detected using the image data when the 3D hexahedral model was projected onto the image, as shown in Figure 24. As such, the outcome of the projected 3D points was expressed in the shape of a trapezoid (near–far effect), for which the distance per pixel was considered to estimate an accurate value. The distance per pixel, p_d, is calculated using Equation (10):

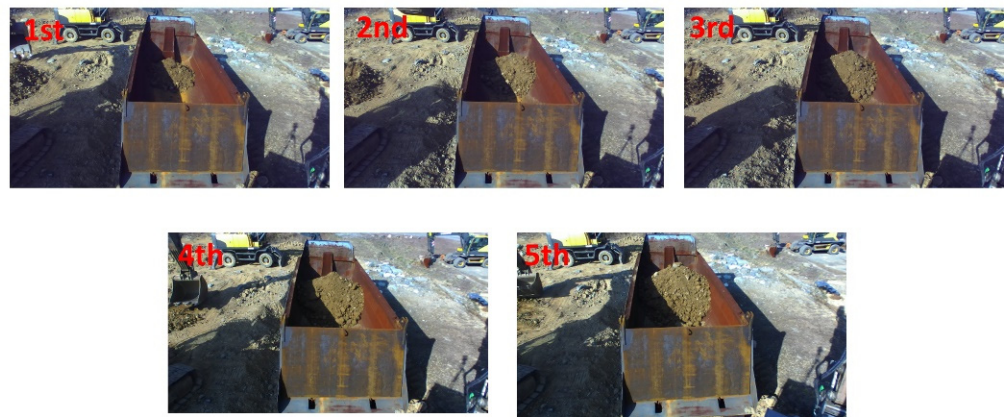$$p_d = real\ distance(mm)\ /\ pixel. \tag{10}$$

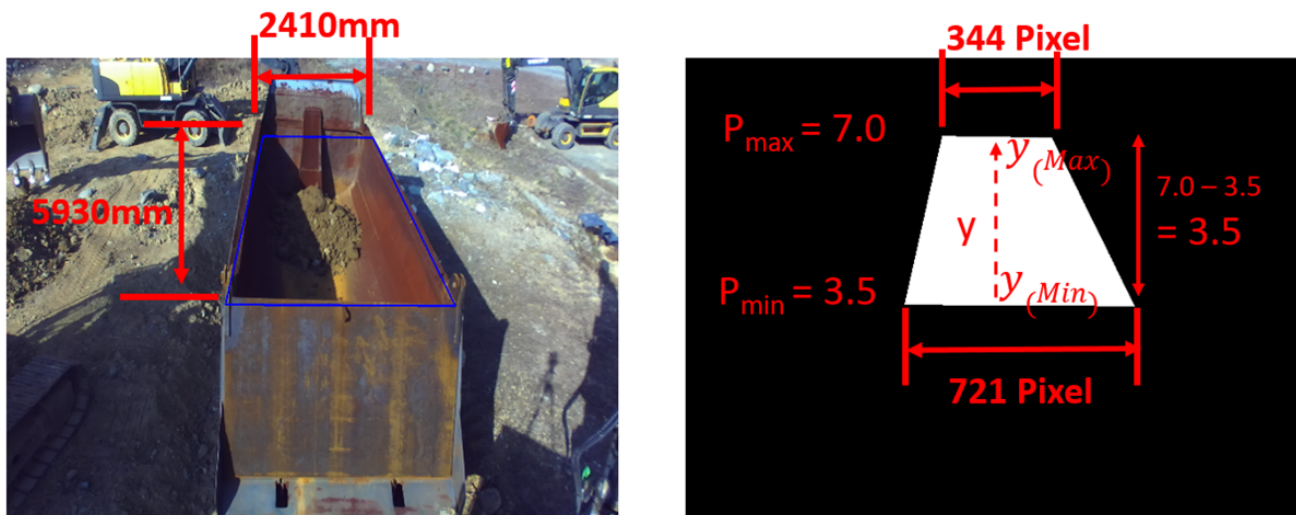**Figure 23.** Volume estimations experiment of load.

Figure 24. Area detection for the part to measure the internal volume of the loading box in the 2D image.

The internal volume of cargo box A obtained from the 2D image is determined using Equations (11)–(13):

$$A = \sum_i \left(P_{i(x)} \times P_{i(y)} \times P_{i(h)} \right), \tag{11}$$

$$P_{i(x)} = P_{min} + (P_{max} - P_{min}) \times \frac{P_{i(y)}}{Y(max)}, \tag{12}$$

$$P_{i(y)} = P_{i(y)}. \tag{13}$$

Here, $P_{i(x)}$ is the height value from the model's floor. The volume of the load was measured using the data from five sets of experiments. The quantitative values were estimated for each loading operation using the difference in volume between frames. The volume of the excavator's bucket was 0.8 m^3; hence, load volume was estimated with the difference in volume estimated from one experiment among the accumulated volume of the load over five sets of experiments.

Table 2 summarizes the measurement results of load volume. Accuracy was determined by the proximity of the value to the ground truth, and the experiment results showed large errors.

Table 2. Volume measurement result of load.

Number of Experiments	Volume Estimation Result (m^3)	Ground Truth (m^3)
2rd	−0.04	0.8
3rd	0.13	1.6
4th	0.13	2.4
5th	1.35	3.2

The sensing device mounted on the excavator was positioned diagonally rather than accommodating a top view. However, as the cargo box is filled with load, more accurate results are possible if the loads are closer to the sensing device and a more precise camera is used.

6. Conclusions

This study applied vision technology that combines 3D LiDAR and stereo distance sensors to estimate a heavy-duty dump truck's cargo box location and volume. The sensing device was installed on the upper section of the excavator's driver seat. First, the plane information for the door of the dump truck's cargo box was extracted, and stereo distance data were combined to obtain the 3D plane data of the rear of the cargo box. Next, a 3D model of the cargo box was generated identically to actual measurements, followed by the determination of the initial 3D position of the box so that the two planes would match. The 3D model was initially adjusted, after which the 3D coordinate value of the longitudinal axis corner was detected iteratively for more precise position estimation. Lastly, the estimated location data were used to convert the stereo distance values into a 3D cargo box model so that the height of the 3D model from the floor surface could be measured. The cargo box interior was then visualized so that loading space data could be delivered to the excavator. Significant errors were detected during the load analysis owing to the distance estimation errors of the stereo camera. Thus, a camera with higher precision is necessary to improve the outcome.

Author Contributions: Conceptualization, J.-h.L., J.L. and S.-Y.P.; methodology, software, validation, formal analysis, J.-h.L.; Sensor Calibration J.L.; investigation, S.-Y.P.; writing—original draft preparation, J.-h.L.; writing—review and editing, S.-Y.P.; supervision, S.-Y.P.; project administration, S.-Y.P.; funding acquisition, S.-Y.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tsuneyama, S.; Takeshita, S.; Tsutsumi, H.; Shirosawa, M. The practical usage of i-construction for enhancing the safety of the construction site. *J. Jpn. Soc. Civ.* **2017**, *73*, I_1–I_6. [CrossRef]
2. Volvo CE Automattion Journey. Available online: https://www.youtube.com/watch?v=FRh2QZaVwoQ (accessed on 2 February 2022).
3. Hyundai Construction Equipment. Available online: http://www.hyundai-ce.com/ko/innovation/ai (accessed on 2 February 2022).
4. Shimano, Y.; Kami, Y.; Shimokaze, K. Development of PC210LCi-10/PC200i-10 machine control hydraulic excavator. *Komatsu Tech. Rep.* **2015**, *60*, 1–7.
5. Vaha, P.K.; Skibniewski, M. Dynamic model of excavator. *J. Aerosp. Eng.* **1993**, *6*, 148–158. [CrossRef]
6. Zhao, J.; Long, P.; Wang, L.; Qian, L.; Lu, F.; Song, X.; Manocha, D.; Zhang, L. AES: Autonomous excavator system for re-al-world and hazardous environments. *arXiv* **2020**, arXiv:2011.04848.
7. Kurinov, I.; Orzechowski, G.; Hämäläinen, P.; Mikkola, A. Automated excavator based on reinforcement learning and multibody system dynamics. *IEEE Access* **2020**, *8*, 213998–214006. [CrossRef]

8. Shariati, H.; Yeraliyev, A.; Terai, B.; Tafazoli, S.; Ramezani, M. Towards autonomous mining via intelligent excavators. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019; pp. 26–32.
9. Schmidt, D.; Proetzsch, M.; Berns, K. Simulation and control of an autonomous bucket excavator for landscaping tasks. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010; pp. 5108–5113.
10. Guan, T.; He, Z.; Manocha, D.; Zhang, L. TTM: Terrain traversability mapping for autonomous excavator navigation in unstructured environments. *arXiv* **2021**, arXiv:2109.06250.
11. Yoo, H.S.; Kim, Y.S. Development of a 3D local terrain modeling system of intelligent excavation robot. *KSCE J. Civ. Eng.* **2017**, *21*, 565–578. [CrossRef]
12. Stentz, A.; Bares, J.; Singh, S.; Rowe, P. A robotic excavator for autonomous truck loading. *Auton. Robot.* **1999**, *7*, 175–186. [CrossRef]
13. Borthwick, J.R. Mining haul truck pose estimation and load profiling using stereo vision. PhD Thesis, University of British Columbia, Columbia, SC, USA, 2009.
14. Yoshida, H.; Yoshimoto, T.; Umino, D.; Mori, N. Practical Full Automation of Excavation and Loading for Hydraulic Excavators in Indoor Environments. In Proceedings of the 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE), Lyon, France, 23–27 August 2021; pp. 2153–2160.
15. Sugasawa, Y.; Chikushi, S.; Komatsu, R.; Kasahara, J.L.; Pathak, S.; Yajima, R.; Asama, H. Visualization of Dump Truck and Excavator in Bird's-eye View by Fisheye Cameras and 3D Range Sensor. In Proceedings of the 16th International Conference on Intelligent Autonomous Systems (IAS), Singapore, 29–31 July 2020; pp. 480–491.
16. Chai, Z.; Sun, Y.; Xiong, Z. Novel method for LiDAR camera calibration by plane fitting. In Proceedings of the 2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Auckland, New Zealand, 9–12 July 2018; pp. 286–291.
17. Fachinotti, V.D.; Anca, A.A.; Cardona, A. A method for the solution of certain problems in least squares. *Int. J. Numer Method Biomed. Eng.* **2011**, *27*, 595–607. [CrossRef]
18. Fischler, M.A.; Bolles, R.C. *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*; ACM: New York, NY, USA, 1981; pp. 381–395.
19. Hirschmuller, H. tereo vision in structured environments by consistent semi-global matching. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; pp. 2386–2393.
20. Konolige, K. Projected texture stereo. In Proceedings of the 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010; pp. 148–155.
21. Desjardins, D.; Payeur, P. Dense stereo range sensing with marching pseudo-random patterns. In Proceedings of the Fourth Canadian Conference on Computer and Robot Vision (CRV'07), Montreal, QC, Canada, 28–30 May 2007; pp. 216–226.
22. Rusinkiewicz, S.; Levoy, M. Efficient variants of the ICP algorithm. In Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, Quebec City, QC, Canada, 28 May–1 June 2001; pp. 145–152.
23. Segal, A.; Haehnel, D.; Thrun, S. Generalized-icp. *Sci. Syst.* **2009**, *2*, 435.
24. Dump Trucks | Volvo Trucks USA. Available online: https://www.volvotrucks.us/trucks/applications/dump/ (accessed on 2 February 2022).
25. Scania. Available online: https://www.scania.com/ (accessed on 2 February 2022).
26. Yoon, J.; Kim, J.; Seo, J.; Suh, S. Spatial factors affecting the loading efficiency of excavators. *Autom. Constr.* **2014**, *48*, 97–106. [CrossRef]
27. Jongluxmanee, J.; Kohei, O.; Yamakita, M. Iterative learning control for soil loading operation of excavator. In Proceedings of the 2019 12th Asian Control Conference (ASCC), Kitakyushu, Japan, 9–12 June 2019; pp. 621–626.