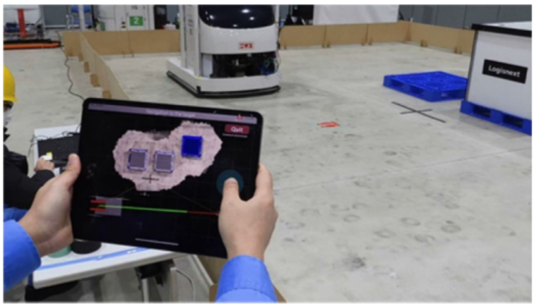
Highly Accurate Navigation Technology for Unmanned Vehicles through Human-Machine Cooperation



YUSUKE KINOUCHI*1 KATSUMASA KITAJIMA*1
NARUHISA KAMEO*1 KIICHI SUGIMOTO*2
YUICHI TAZAKI*3 YASUYOSHI YOKOKOHJI*4

In order to increase the flexibility of our unmanned vehicle products such as automated guided forklifts (hereinafter, AGF(s)) and plant patrol inspection robots, and promote further application of them at sites, a navigation technology that allows accurate guidance of an unmanned vehicle to an object with unclear position and posture is necessary. Therefore, Mitsubishi Heavy Industries Ltd. (MHI) has developed, jointly with Kobe University, an accurate navigation technology through human-machine cooperation that allows positioning of a target object in cooperation with a tablet terminal carried by a worker and unmanned vehicle sensor and accurate guidance of the unmanned vehicle based on the positioning information. We plan to apply this technology to contribute to the increasing intelligence and sophistication of the mobility domain in MHI group business.

1. Introduction

In a place like a temporary storage area for cargo, which can be seen in certain sections of a plant or warehouse, objects are placed anywhere at the site worker's discretion and it is difficult to introduce an unmanned vehicle. In such an environment, the number of objects and their positions/postures in the section are indefinite and there are many blind areas for the unmanned vehicle sensors depending on the locations of the objects, and it is difficult to make the unmanned vehicle recognize the location of a target object with high accuracy and to accurately guide the vehicle. Therefore, we have developed an accurate navigation technology through human-machine cooperation that allows positioning of a target object in cooperation with a tablet terminal carried by a worker (hereinafter, "worker's terminal") and unmanned vehicle sensors and accurate guidance of the unmanned vehicle based on the positioning information, as shown in **Figure 1**.

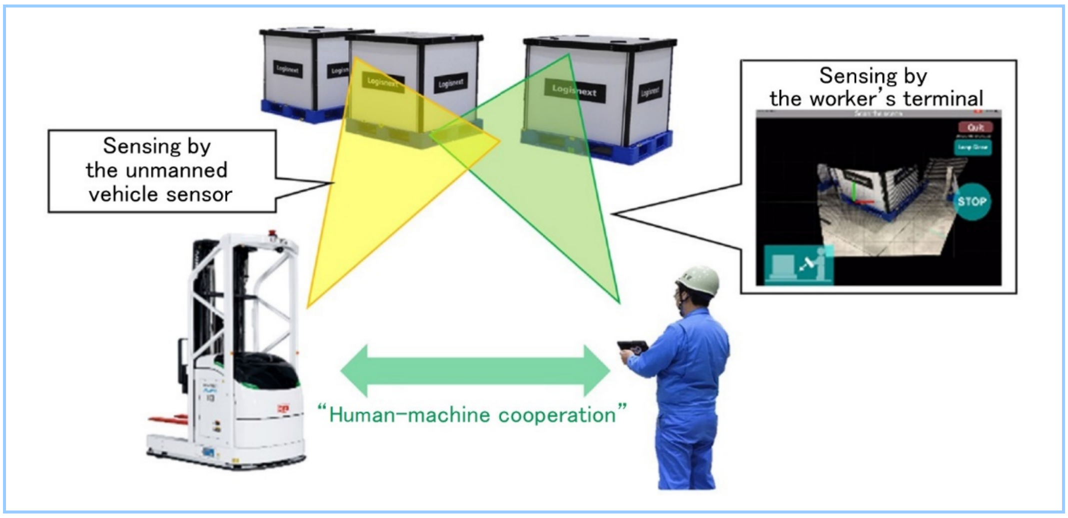


Figure 1 Image of application of the human-machine cooperation technology for AGFs

*1 CIS Department, Digital Innovation Headquarters, Mitsubishi Heavy Industries, Ltd.
*2 Senior Manager, CIS Department, Digital Innovation Headquarters, Mitsubishi Heavy Industries, Ltd.
*3 Associate Professor, Department of Mechanical Engineering, Graduate School of Engineering, Kobe University
*4 Professor, Department of Mechanical Engineering, Graduate School of Engineering, Kobe University

For application of this technology, it is not necessary to additionally install any special sensor in the unmanned vehicle and also any infrastructure sensor such as a surveillance camera, and the system can be constructed at low cost. In addition, unlike the case of the conventional remote operation of an unmanned vehicle with human direct control, this technology supports the autonomous behavior of an unmanned vehicle with minimal human intervention required. Thus, this technology allows anyone to handle an unmanned vehicle easily without any special operation skills. Furthermore, it eliminates the necessity of a worker having to operate the unmanned vehicle for a long time and can provide a solution to reducing on-site workload.

This report describes the overview of the developed human-machine cooperation technology and the result of the effectiveness evaluation for the prototype system constructed using an actual AGF.

2. Human-machine cooperation technology

2.1 Human-machine cooperation sensing and navigation

In the developed sensing technology in cooperation of a worker's terminal and unmanned vehicle sensors, it is based on the premise that the worker's terminal sensor (3D LiDAR) can scan the surrounding environment and obtain three-dimensional point cloud data (hereinafter, the "worker's terminal map") which represents the three-dimensional shape of the target object as a collection of points in a three-dimensional space. It is also premised that the unmanned vehicle sensors are 2D LiDAR for collision avoidance which is installed in the underbody of an industrial vehicle such as an AGF, and can obtain two-dimensional point cloud data (hereinafter, the "unmanned vehicle's map") which represents the outline shape of the cross-section of the target object cut at the scan plane as a collection of points in a two-dimensional plane.

As described in Chapter 1, in an environment where objects are placed anywhere by human discretion, it is difficult to make the unmanned vehicle recognize the location of a target object with high accuracy. Especially, depending on the arrangement of objects, blind areas of the unmanned vehicle sensors may develop and the unmanned vehicle cannot recognize any object in the blind areas. Therefore, we have developed the human-machine cooperation sensing technology which allows the unmanned vehicle to recognize the positions/postures of the objects that cannot be directly observed by the unmanned vehicle sensors by complementing the blind areas of the unmanned vehicle sensors by the sensing result of the worker's terminal.

As shown in **Figure 2**, when the developed technology is applied, the blind areas of the unmanned vehicle sensors can be "complemented" by the worker's terminal map. Specifically, the unmanned vehicle map and the worker's terminal map individually detect objects, and the objects detection result of the worker's terminal map is substituted for information of the objects that cannot be detected by the unmanned vehicle map (objects in the blind areas). However, in order to realize this method, it is necessary to obtain coordinate transformation accurately so that the unmanned vehicle map and the worker's terminal map are superposed.

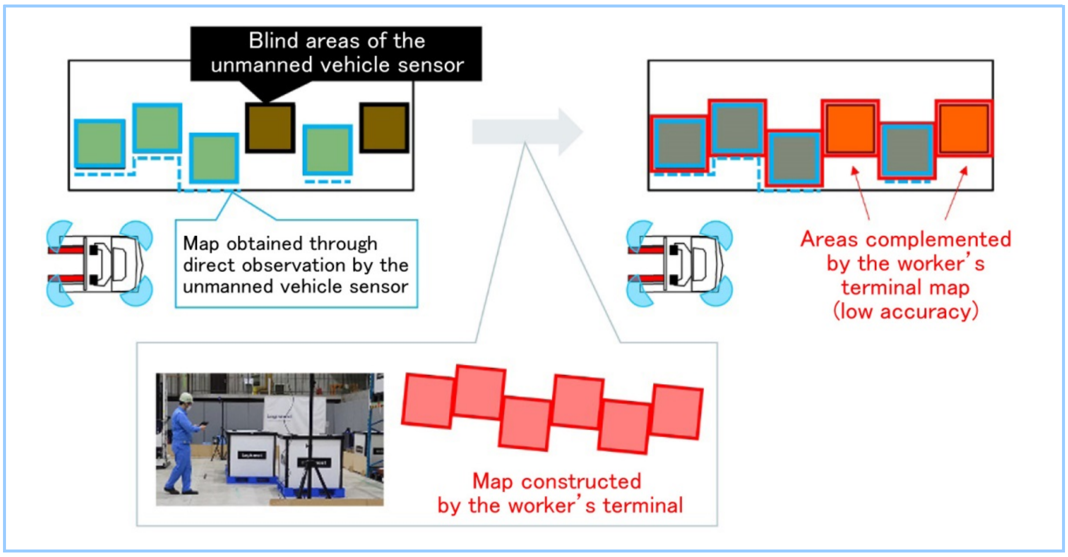


Figure 2 Human-machine cooperation sensing

Even if blind areas of the unmanned vehicle sensors can be complemented by the worker's terminal map, there is a limit to the positioning accuracy of the worker's terminal which is not for industrial use, and it is difficult to estimate the positions/postures of target objects with the accuracy required for accurate guidance of the unmanned vehicle. Therefore, in the developed navigation technology, the map information of the worker's terminal is treated as an area to be used as reference (hereinafter, "reference area") for implementation of more accurate sensing by the unmanned vehicle sensors so that errors in the worker's terminal map do not directly affect the guidance accuracy of the unmanned vehicle. Specifically, using the reference area by the worker's terminal as a clue, the unmanned vehicle autonomously moves to the vicinity of a target object. When the unmanned vehicle reached the point where the unmanned vehicle sensors can directly observe the inside of the reference area, high-accuracy positioning of the target object is performed again, and based on the positioning result, a final approach path to the target object is planned.

2.2 Map merging

(1) Overview

In order that the blind areas of the unmanned vehicle sensors can be complemented by the worker's terminal map, we have developed a map merging technology for obtaining coordinate transformation that allows the unmanned vehicle map and the worker's terminal map to be accurately superposed.

The basic idea for map merging is that overlap regions of an unmanned vehicle map and a worker's terminal map are found and the coordinate transformation by which the overlap regions coincide with each other as much as possible is inversely estimated by optimization calculation. This time, however, the unmanned vehicle and the worker's terminal greatly differ in the sensor's view point, dimension of measured data, distance measurement characteristics, etc. So, when the raw data of the maps is used, it is difficult to find overlap regions of the maps and to perform calculations for making the overlap regions coincide with each other accurately.

Therefore, in the development of this technology, we used an object having a known shape in each map as a reference object for map merging (**Figure 3**). Assuming that this technology was applied to logistics sites, we used a pallet having a standardized shape as a reference object. With the use of a reference object, it becomes possible to obtain coordinate transformation for robustly superposing the maps in spite of the difference in sensor between the unmanned vehicle and the worker's terminal.

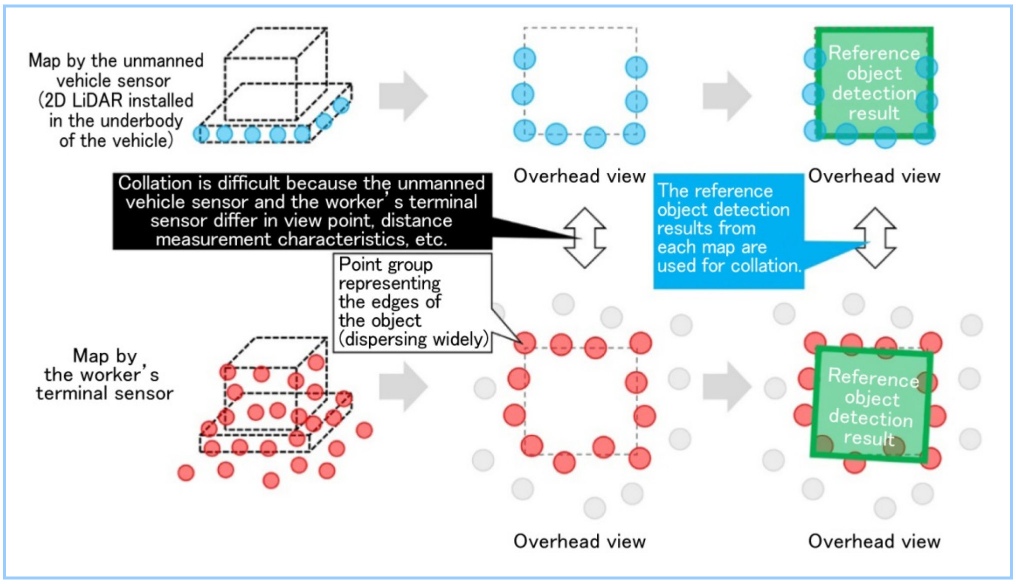


Figure 3 Use of reference object in map merging

(2) Processing procedure for map merging

The processing procedure of the map merging technology using a reference object (pallet) is described below.

(i) Detection of pallets

Pallets are detected from each of the worker's terminal map (three-dimensional point cloud) and the unmanned vehicle map (two-dimensional point cloud) by point cloud processing. In the palette detection from the worker's terminal map, the proximity points(1) were utilized. The proximity points are the points in the point cloud observed by the sensor where the distance to the observation point (sensor origin) is minimal (the smallest in the proximity point group). By using the proximity points, it is possible to extract many point clouds at the corners of the pallet, which are important in pallet detection.

(ii) Generation of pallet maps

Two-dimensional location data of the pallets (hereinafter, "pallet map") is generated from each of the worker's terminal map and the unmanned vehicle map.

(iii) Collation of pallet maps

Based on the initial value $T_0 \in \mathbb{R}^{3\times3}$ for map coordinate transformation, coordinate transformation of the pallet map of the worker's terminal is performed.

(iv) Merging of pallet maps

Coordinate transformation $T \in \mathbb{R}^{3\times3}$ by which errors in positions/postures between the linked pallets are minimized is derived by optimization calculation. The result of the coordinate transformation T of the pallet map of the worker's terminal is used to complement the missing parts on the pallet map of the unmanned vehicle.

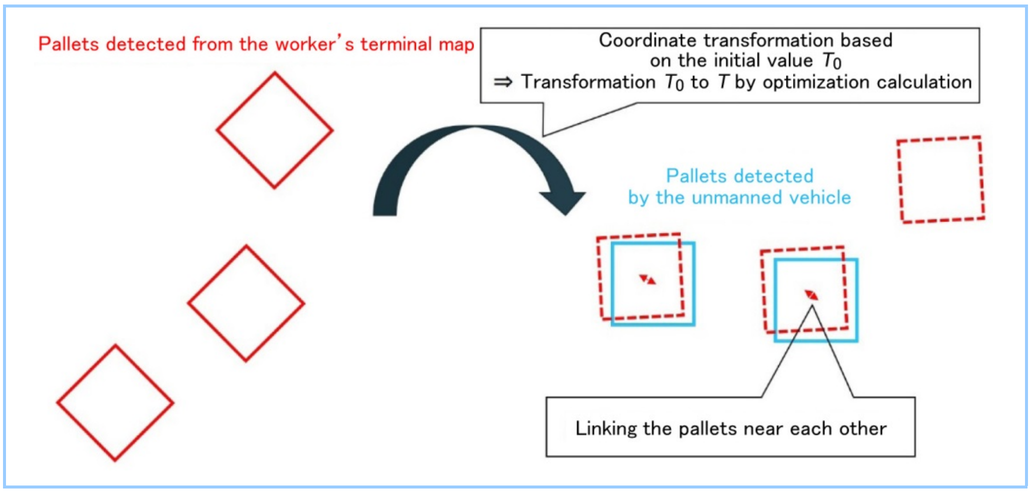


Figure 4 Collation with reference object

(3) Optimization calculation method

The method for deriving the coordinate transformation T by optimization calculation in the above procedure (iv) is described. When the position and posture of the reference object detected from the unmanned vehicle map are q_i, ψ_i $(i = 1, \ldots, N)$, respectively, and the position and posture of the reference object detected from the worker's terminal map are p_i, ϕ_i $(i = 1, \ldots, N)$, respectively, T is obtained by the following optimization calculation (w is the weighting factor of error in position/posture):

$$T = \underset{T}{\operatorname{argmin}}\, E \quad \text{where } E = \frac{1}{2}\sum_i \left(\|Tq_i - p_i\|_2^2 + w(\theta + \psi_i - \phi_i)^2\right) \#(1)$$

When $T = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix}$ (homogeneous transformation matrix of two-dimensional plane), x, y and θ that minimize the evaluation function E only need to be obtained. At the point where the evaluation function is minimized, the partial differential becomes zero. Therefore, the simultaneous equation that should be satisfied by the solution to be obtained can be set up as follows, where "$\bar{\cdot}$" represents the average.

$$\frac{\partial E}{\partial x} = 0, \quad \frac{\partial E}{\partial y} = 0, \quad \frac{\partial E}{\partial \theta} = 0 \iff \begin{cases} x = -\overline{q_x}\cos\theta + \overline{q_y}\sin\theta + \overline{p_x} \\ y = -\overline{q_x}\sin\theta - \overline{q_y}\cos\theta + \overline{p_y} \\ f(\theta) = 0 \end{cases}$$

$$\text{where} \begin{cases} f(\theta) = \sqrt{\alpha^2+\beta^2}\sin\left(\theta + \arctan\dfrac{\alpha}{\beta}\right) + \dfrac{w}{N}\sum_i(\psi_i - \phi_i) + w\theta \\ \alpha = -\overline{p_x}\,\overline{q_y} + \dfrac{1}{N}\sum_i p_{ix}q_{iy} + \overline{p_y}\,\overline{q_x} - \dfrac{1}{N}\sum_i p_{iy}q_{ix} \\ \beta = -\overline{p_x}\,\overline{q_x} + \dfrac{1}{N}\sum_i p_{ix}q_{ix} - \overline{p_y}\,\overline{q_y} + \dfrac{1}{N}\sum_i p_{iy}q_{iy} \end{cases} \qquad \#(2)$$

$f(\theta) = 0$ is numerically determined by the Newton-Raphson method. When the result is substituted into the above equation, x, y can be obtained. Thus, T is derived.

(4) Method for estimating the initial value T_0 for map coordinate transformation

T_0 in the above procedure (iii) only needs to be estimated with some degree of accuracy and there are various possible methods. In the prototype system described below, a process of first taking a photo of the marker on the floor by the worker's terminal was implemented. The unmanned vehicle has estimated its self-location and the origin point of the map obtained by the unmanned vehicle sensor is already known. Therefore, when the origin point of the map by the worker's terminal is estimated by the marker or the like, T_0 can be obtained.

3. Effectiveness evaluation using the prototype system

3.1 Prototype system

Figure 5 shows the human-machine cooperation flow with the prototype system. After the worker's terminal is started, when the origin point of the map (marker on the floor) is captured at the center of the field angle of the worker's terminal camera according to the instruction on the screen, the initial value T_0 for map coordinate transformation to be used in the map merging process is obtained. After the origin point of the map is photographed, the inside of the section including the target object is scanned by the worker's terminal and the map is constructed.

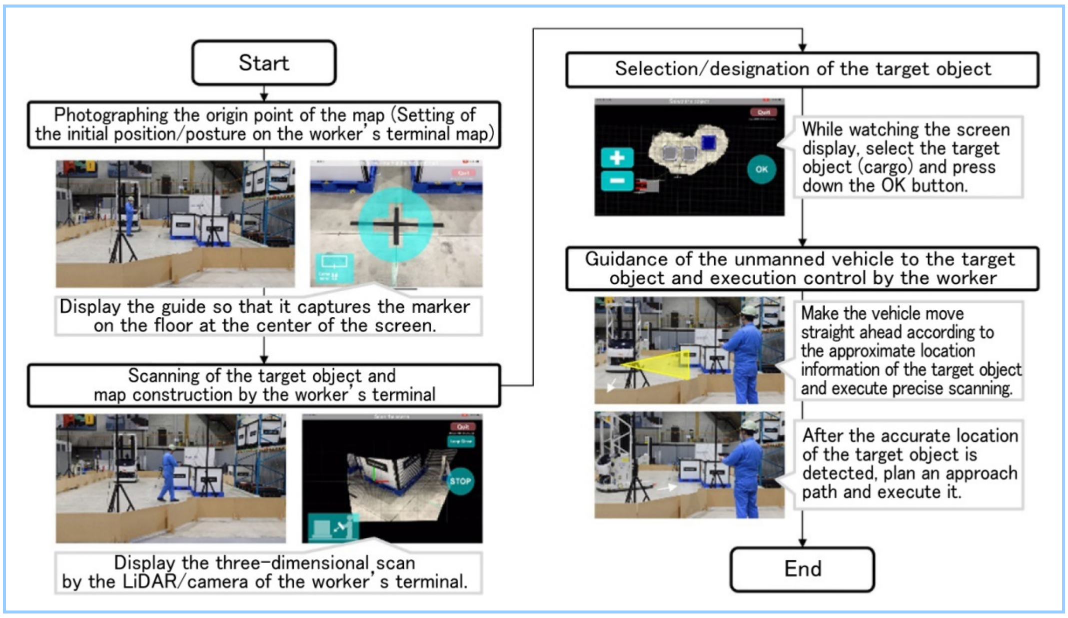


Figure 5 Human-machine cooperation flow of the prototype system

After the map is constructed by the worker's terminal, designation of the target object and guidance of the unmanned vehicle are performed using the user interface (UI) shown in **Figure 6**. On the "Select the object" screen, the map obtained by the worker's terminal is displayed in an overhead view and the user selects the target object, to which the unmanned vehicle is guided, on the map display. Next, when the screen moves to the "Guidance" screen, the guide path for the unmanned vehicle is displayed. The unmanned vehicle moves along the path while the user keeps pressing down the button, and when the user releases the button, the unmanned vehicle temporarily stops.

As described above, this system has a UI designed to allow intuitive instructions to the unmanned vehicle and even a user with no expertise or special operation skills regarding the unmanned vehicle can easily handle it.

This human-machine cooperation technology has been developed as a module of ROS (Robot Operating System)[2] which is a robot software development framework and it can be easily added to existing unmanned vehicle systems constructed by ROS. Actually, this prototype system was constructed by system integration of the developed human-machine cooperation technology and MHI's autonomous mobile robot technologies[3][4] on ROS. For many functions of this prototype system described in Section 2.1, such as high-accuracy positioning of a target object in the reference area and planning of the approach path to the target object, the technologies already developed with ROS were utilized for improving system development efficiency.

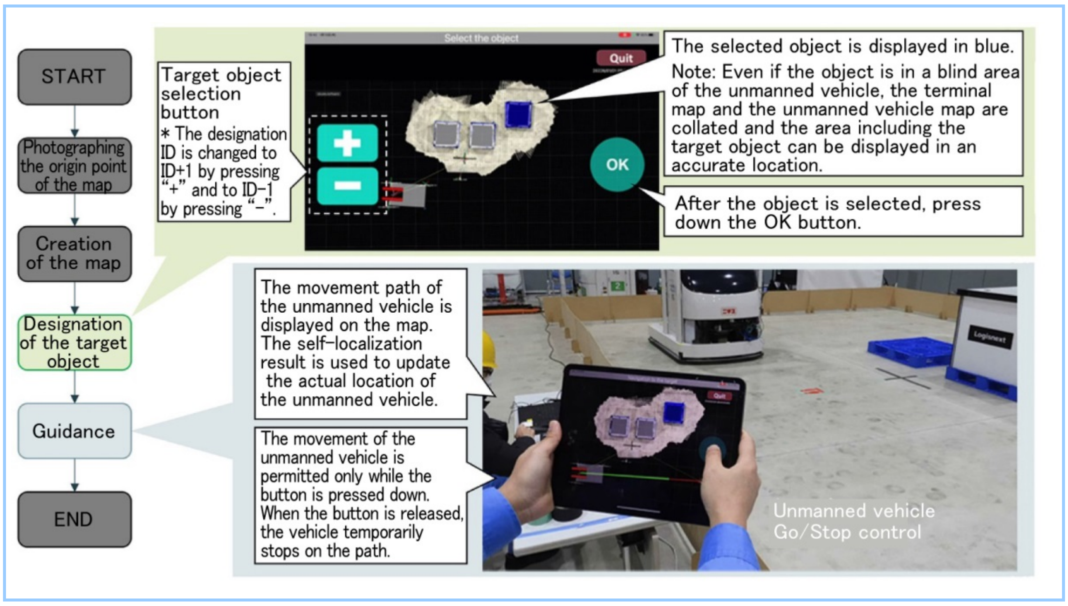


Figure 6 User interface for the prototype system

3.2 Effectiveness evaluation

In order to evaluate the effectiveness of the prototype system, we conducted a test with actual equipment. **Figure 7** shows the test area and the test vehicle. The object in this test is a cargo loaded on a pallet. To evaluate the accuracy of guidance to the target object, we set cameras for motion capture in the test area and set the target markers on the test vehicle and the target object. It should be noted that the cameras for motion capture and the markers were set for this test and they are not needed in the actual operation of this technology.

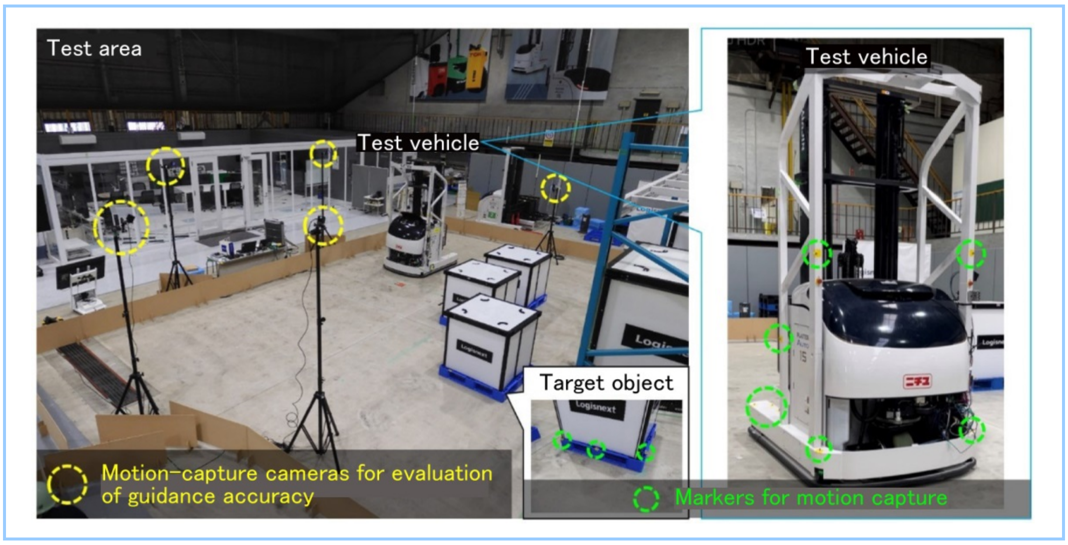


Figure 7 Test area and test vehicle

Figure 8 shows the content of the test. In this test, the construction of a map by the worker's terminal and the guidance of the unmanned vehicle to the target object in the blind area of the unmanned vehicle sensors using the map were repeatedly conducted three times.

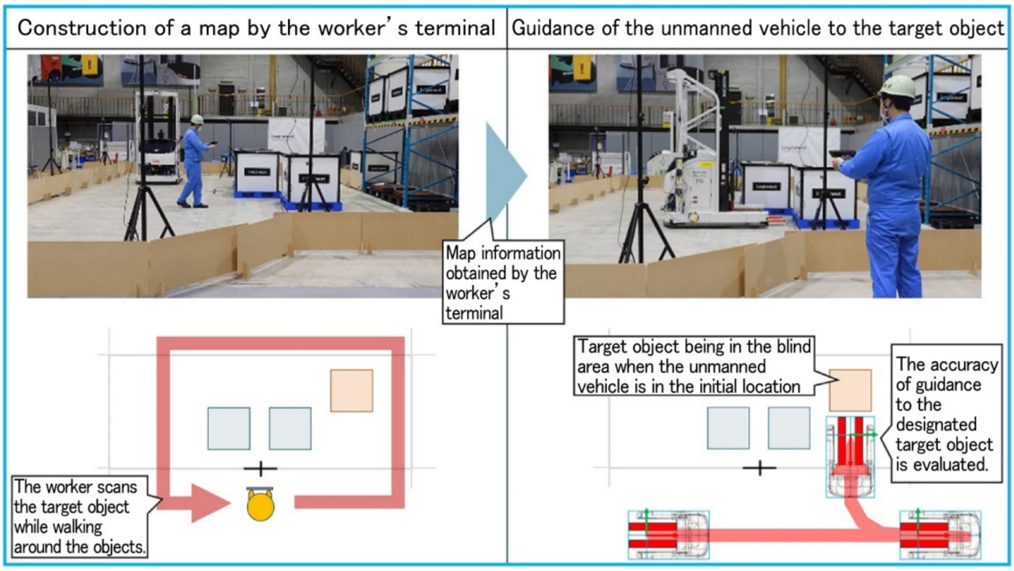


Figure 8　Content of the test

Figure 9, **Figure 10** and **Figure 11** show the test results. The green frames in Figure 9 show the locations of the target objects recognized by the unmanned vehicle on the merged map. As shown in Figure 9, in all three tests, the target object in the blind area of the unmanned vehicle sensors were recognized by the unmanned vehicle. However, as shown in the center of Figure 9 ($n = 2$), we also confirmed cases where the error of the detection position of the object is large. As described in Section 2.1, there is a limit to the accuracy of detecting an object by a worker's terminal, so it is necessary to guide an unmanned vehicle robustly against this error.

Figure 10 shows the scene where the unmanned vehicle is guided to the target object, and it was confirmed that the vehicle could be guided to the pallet in the blind area of the unmanned vehicle sensors with the accuracy that allowed the forks of an AGF to be inserted into the insertion part of the pallet.

Figure 11 shows the result of the quantitative evaluation for the accuracy of guidance to the target object in the blind area of the unmanned vehicle sensors using the motion-capture measured values. As shown in Figure 11, in three tests, the maximum translation error was 10.3 mm and the maximum angle error was 0.35 deg. Thus, it was confirmed that the developed technology allowed accurate guidance of an unmanned vehicle to a target object. Especially, $n = 2$ in Figure 11 shows a case in which an unmanned vehicle was guided by using the result that the accuracy of detecting an object by a worker's terminal was the lowest (the center of Figure 9). It was confirmed that the unmanned vehicle could be guided precisely without significant effects compared to the other two cases ($n = 1,3$).

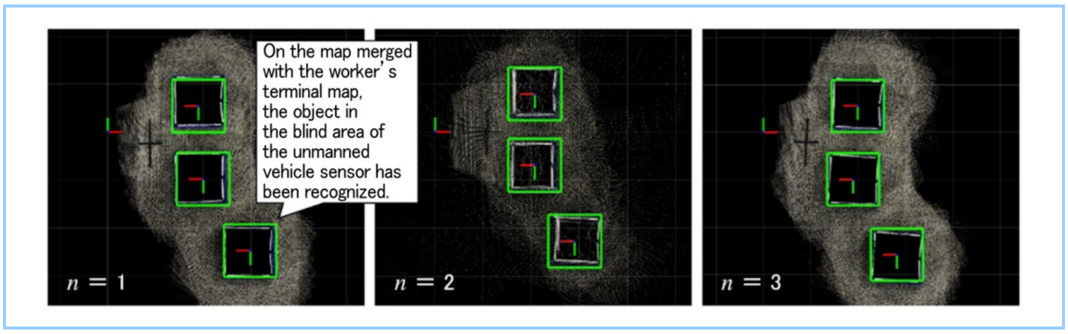


Figure 9　Target object recognition result on the merged map

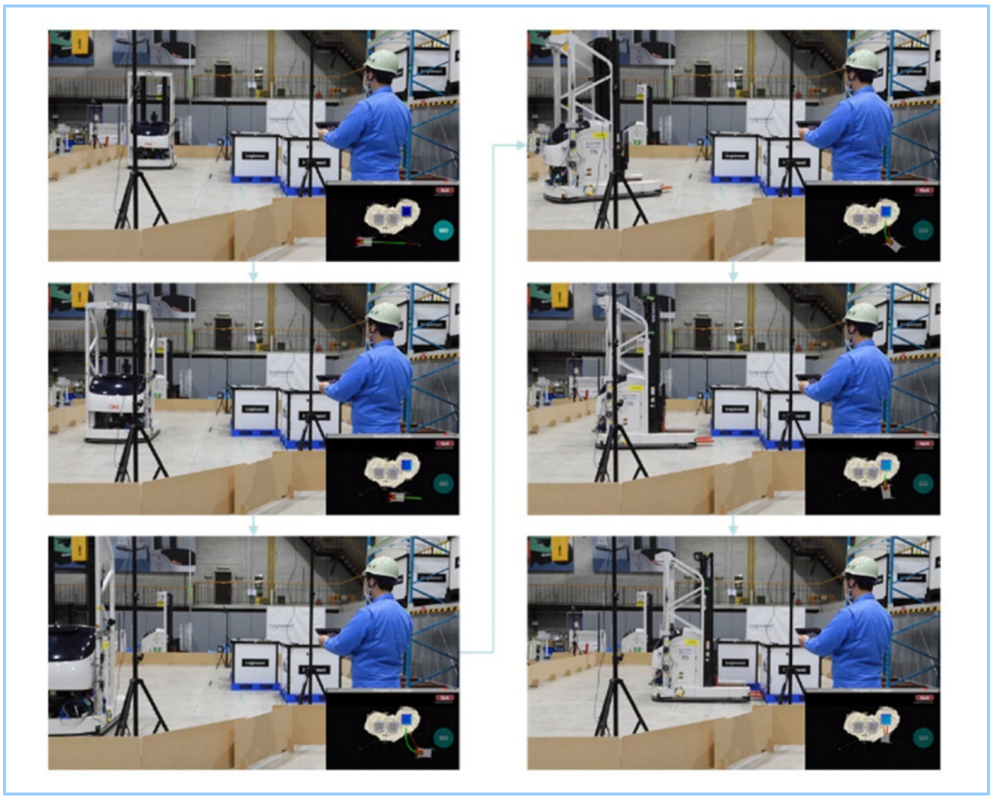


Figure 10 Guiding the unmanned vehicle to the target object

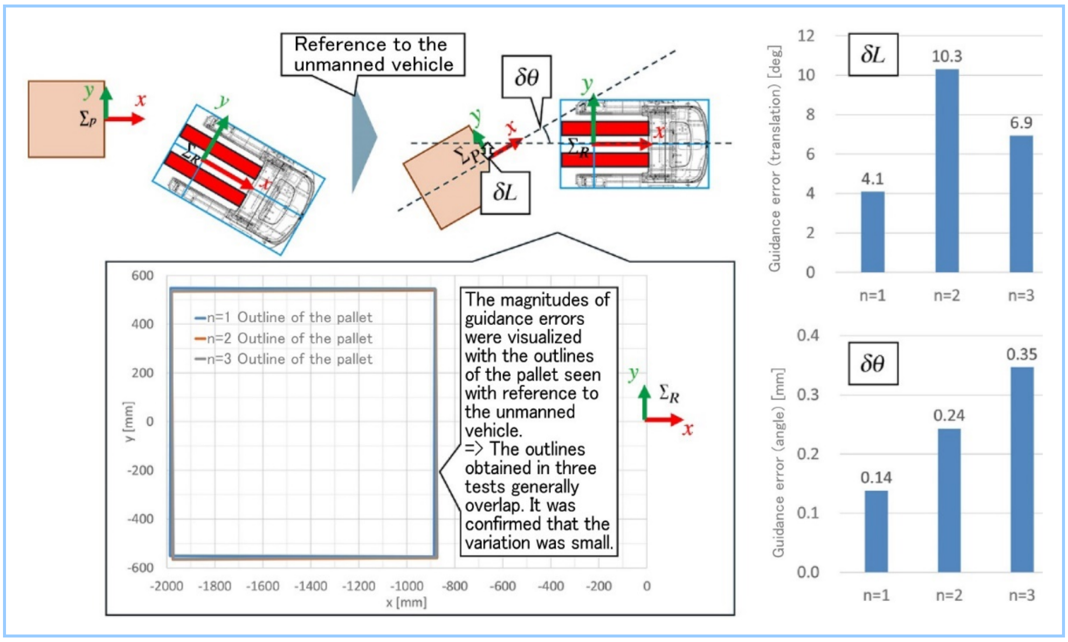


Figure 11 Evaluation result for the accuracy of guidance to the target object in blind area of the unmanned vehicle sensor

4. Conclusion

In order to improve the flexibility of our unmanned vehicle products and promote further application to sites, we developed the accurate navigation technology for unmanned vehicles through human-machine cooperation. The application of this technology allows accurate guidance of an unmanned vehicle to a target object with minimal human intervention required even in an environment like a temporary storage area for cargo in a logistics site where the number of objects and their positions/postures are indefinite and there are many blind areas of the unmanned vehicle sensors. From the evaluation result for the effectiveness of the prototype system constructed using an actual AGF, we confirmed that it could guide an unmanned vehicle with sufficient accuracy to a target object in the blind area of the unmanned vehicle sensors.

Going forward, we will apply this technology to our unmanned vehicle products which require smooth operation in cooperation with site workers. We will also further develop this technology and promote research and contribute to the increasing intelligence and sophistication of the mobility domain in MHI group business.

References

(1) Tazaki, Y. et al., Outdoor Autonomous Navigation Utilizing Proximity Points of 3D Pointcloud, Journal of Robotics and Mechatronics, Vol.32 No.6 (2020) p.1183-1192
(2) Quigley, M. et al., ROS: an open-source Robot Operating System, ICRA. (2009)
(3) Yusuke Kinouchi et al., Robots utilizing ROS/Gazebo in Mitsubishi Heavy Industries, Journal of the Robotics Society of Japan, Vol.35 No.4 (2017), p.8-11
(4) Atsushi Wada et al., Feasibility Study of Pallet Handling in Mixed Fleet Environment, Mitsubishi Heavy Industries Technical Review, Vol.58 No.2 (2021)