2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 1-4, 2023, Oahu, Hawaii, USA

Development of a Calibration Method of Hidden Background Observer Cameras for Diminished Reality Using Radio Direction Finding

Tasuku Ono[1], Kimi Ueda[1], Hirotake Ishii[1], Hiroshi Shimoda[1]

Abstract— **Hidden background cameras used to realize Diminished Reality (DR), which allows the observer to see through obstacles, are installed behind the obstacles. In order to realize DR, it is necessary to measure the positions of these cameras with respect to the observer's position, but this is difficult because the cameras cannot be seen directly from the observer's position. In this study, two calibration methods, the intersection method and the P4P method, which use the directions of radio waves from radio sources mounted on each camera to measure the positions of the hidden background cameras, are proposed. The intersection method measures the direction from several different points to the hidden background cameras and estimate the 3D position of the hidden background camera as the intersection of the measured directions. The P4P method measures the relative pose between multiple hidden background cameras in advance and uses it as a constraint to estimate the 3D poses of the hidden background cameras. The results showed that the P4P method was more accurate than the intersection method, and able to calibrate the camera with a position error of less than 100 mm.**

I. Introduction

By using Diminished Reality (DR) to remove obstacles in images, it is possible to see what is behind them. A typical system to realize such DR usually consists of hidden background cameras installed behind obstacles and a tablet PC with a camera that presents DR experience to the user. Images captured by the hidden background camera is processed and projected onto the image captured by the camera held by the user [1]. In such a system, it is necessary to measure the position and orientation (hereafter, pose) of the hidden background cameras with respect to the user's pose in real time. However, the hidden background cameras are behind obstacles and cannot be directly seen from the user, making it difficult to determine the poses of the hidden background cameras.

On the other hand, there is a technique to measure the direction in which the radio source exists by receiving a radio wave with an antenna. Such radio waves can go through walls and other obstacles, as long as the walls do not contain large metal objects. Therefore, it is considered possible to measure the poses of the hidden background cameras through the obstacles by mounting radio wave transmitters on the hidden background cameras. However, if such directional measuring technique is used as is, the accuracy will not be enough for DR because the measurement includes much error. The goal of this study, therefore, is to develop a new technique to estimate the hidden background camera poses using radio direction measurement with errors.

*This work was not supported by any organization

[1]Tasuku Ono, Kimi Ueda, Hirotake Ishii and Hiroshi Shimoda are with Graduate School of Energy Science, Kyoto University, Yoshida-honmachi, Sakyo, Kyoto, 606-8501 Japan {tasuku, ueda, hirotake, shimoda}@ei.energy.kyoto-u.ac.jp

II. Related works

There are two main algorithms for position estimation: range-based and range-free. The range-based method measures distance and angle when estimating position. The measurement methods include angle of arrival (AoA), time of arrival (ToA) [2], time difference of arrival (TDoA) [3], received signal strength (RSS) [4]. ToA and TDoA require facilities to transmit radio waves at the nodes, and accurate time synchronization between the node and antenna. RSS cannot be used in environments with high time-varying characteristics. In addition, these studies assume an environment where the presumed nodes, not the antennas, move around, but mounting antennas on each of the hidden background cameras would be larger in size and more expensive than mounting transmitters.

Range-free methods utilize only network connectivity, and include the centroid algolithm, DV-HOP, and APIT [5]. Since these methods are inaccurate, they are not expected to be used for pose estimation of hidden background cameras, which requires accuracy.

Furthermore, Tara et al. [6] presented a method to estimate the location of a tag using SAR by moving with an antenna, but this method requires a dedicated high-precision antenna.

III. Proposed methods

This section describes two hidden background camera calibration methods proposed in this study.

A. Overview

The proposed system configuration is shown in Fig. 1. In this study, we assume a situation in which multiple hidden background cameras are positioned behind an obstacle, such as an opaque wall. The goal of the proposed method is to estimate poses of the hidden background cameras with reference to a specific point in front of the obstacle without directly seeing the behind of the obstacle. At this time, the target environment is assumed to be static. Each hidden background camera is equipped with a tag, which is a radio source used for direction detection and identification. For simplicity, we will refer to a hidden background camera and a tag as "a hidden background device," and assume that the tag and the camera are located in the same position.

The calibration methods are outlined below (Common outlines for the two proposed methods.) First, (Step 1) hidden

2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 979-8-3503-3702-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/SMC53992.2023.10394605

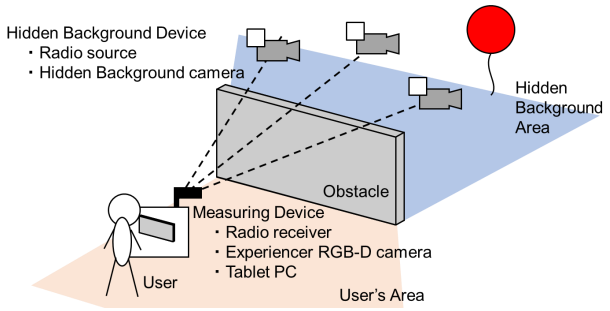


Fig. 1. System configuration for calibration by radio wave

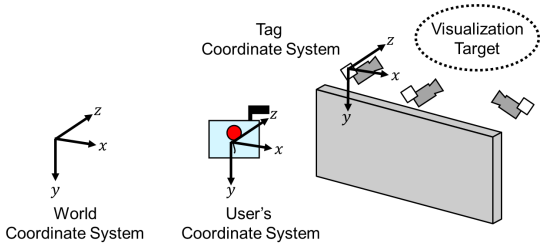


Fig. 2. Coordinate systems to be considered

background devices equipped with a wireless radio source are installed in the area behind the obstacle (henceforth, hidden background area), and their relative 3D poses are measured in advance (There are no obstacles between the hidden background cameras, so the relative pose between them is relatively easy to measure.) Next, (Step 2) move around with the RGB-D camera and antenna (hereafter, a measuring device) in the area where the user moves when using DR, capture images of the user's area, and measure the direction in which the hidden background devices exist through the visually opaque obstacle. Then, (Step 3) the captured images are processed using structure from motion (SfM) technique to obtain the pose of the measuring device at the time each image was captured. After that, (Step 4) integrate multiple data measured at close locations to obtain a value with reduced error in the measurement results. Finally, (Step 5) calculate the poses of the hidden background devices using the error-reduced data.

B. Definition of coordinate systems

The coordinate systems used in this study are shown in Fig. 2. All coordinate systems are right-handed. The origin of the world coordinate system is the focal point of the user's RGB-D camera when the user's area is first captured, the shooting direction of the camera at that time is the positive direction of the z-axis, and the horizontal right direction of the user's camera image is the positive direction of the x-axis. The origin of the tag coordinate system is the position of one of the hidden background devices, the shooting direction of the hidden background camera is the positive direction of the z-axis, and the horizontal right direction of the hidden background camera image is the positive direction of the x-axis. The origin of the user's coordinate system is the focal point of the camera of the measuring device, the camera's shooting direction is the positive direction of the z axis, and the horizontal right direction of the camera image is the positive direction of the x axis.

C. The intersection method

In this subsection, we describe the first hidden background camera calibration method proposed in this study. In this method, the positions of the hidden background devices are determined from the intersection of vectors in the direction in which the devices are located (hereafter, the intersection method.) The premise for using this method is that three or more hidden background devices are placed behind the obstacle at a certain distance from each other, and that there is a large user's area in front of the obstacle where the user can move around with the measuring device. All hidden background cameras shall be pointed in a direction that allows them to capture a visualization target space simultaneously. The antenna used for measurement shall be capable of measuring the azimuth angle and elevation angle in the direction in which the tag emitting radio waves exists.

As the Step 1 in the intersection method, first place three or more hidden background devices at a sufficient distance from each other so that at least two cameras can simultaneously capture about one meter square. Next, the relative poses between the hidden background devices are measured by a checkerboard-based calibration method [7] [8].

Next, as the Step 2, the measuring device is used to take pictures while moving around the user's area. At this time, the direction of the tag is also recorded simultaneously using the antenna.

Next, as the Step 3, the captured images are processed using SfM technique to determine the camera pose at the time each image was acquired.

The measurement results for the directions in which hidden background devices are located contain errors. Therefore, as the Step 4, the multiple tag direction data measured in close proximity are averaged to reduce error. In this study, we attempt to reduce error by dividing the space where the measuring device was moved during the measurement into equally spaced grids and accumulating the directional information for each grid. First, the area over which the measuring device was moved when the user's area was measured is divided into equally spaced grids. Then, using the position of the measuring device obtained by SfM, we find out which grid the device was in at the moment of each measurement, and find the directional information obtained from the measurement within that grid. The directional unit vector $\boldsymbol{u}_{ijt}$ from the camera to the hidden background device j in the world coordinate system is expressed as

$$\boldsymbol{u}_{ijt} = \boldsymbol{M}_t(\cos\theta_{jt}\sin\phi_{jt}, -\sin\theta_{jt}, \cos\theta_{jt}\cos\phi_{jt}, 1)^\top \tag{1}$$

where $\boldsymbol{M}_t$ is the pose of the measuring device at time t in the world coordinate system expressed as the rotation matrix in a homogeneous coordinate system, ϕ_{jt} is the measured Azimuth angle, and θ_{jt} is the Elevation angle.

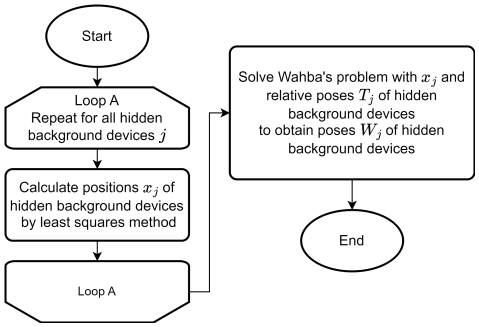


Fig. 3. Flowchart of pose estimation of hidden background devices in the intersection method

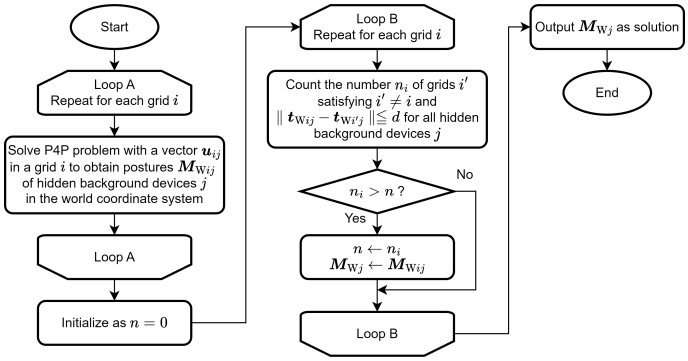


Fig. 4. Flowchart of pose estimation of hidden background devices in the P4P method

The vector $\boldsymbol{u}_{ij}$ is obtained by averaging this $\boldsymbol{u}_{ijt}$ for each grid to reduce the error. The larger the number of $\boldsymbol{u}_{ijt}$ to be added, the more reliable the result is considered to be.

Finally, as the Step 5, the average of the direction vectors obtained in Step 4 is used to estimate the poses of the hidden background devices. The process flow of the pose estimation is shown in Fig. 3.

First, we find the positions $\boldsymbol{t}_{\mathrm{W}j}$ in the world coordinate system for all hidden background devices j. The point at which the sum of the squares of the distances from each straight line is the smallest is calculated using the least squares method.

Then, from the position $\boldsymbol{t}_{\mathrm{W}j}$ of the hidden background device j in the world coordinate system and the pose $\boldsymbol{M}_{\mathrm{T}j}$ of the hidden background device in the tag coordinate system, Wahba's problem [9] is solved to obtain the pose $\boldsymbol{M}_{\mathrm{W}j}$ of the hidden background device in the world coordinate system. Specifically, first, the vectors from the centers of gravity of all the hidden background devices to the position of each hidden background device in the world coordinate system and tag coordinate system $\{\boldsymbol{m}_j\}, \{\boldsymbol{m'}_j\}$ are calculated with $\boldsymbol{m}_j = \boldsymbol{t}_{\mathrm{W}j} - \bar{\boldsymbol{t}_{\mathrm{W}}}, \boldsymbol{m'}_j = \boldsymbol{t}_{\mathrm{T}j} - \bar{\boldsymbol{t}_{\mathrm{T}}}$. Here, $\boldsymbol{t}_{\mathrm{T}}$ and $\boldsymbol{t}_{\mathrm{W}}$ are the centers of gravity of all the hidden background devices in the tag coordinate system and the world coordinate system respectively.

The 3×3 rotation matrix $\boldsymbol{R}$ satisfying

$$\sum_j \| \boldsymbol{m}_j - \boldsymbol{R}\boldsymbol{m'}_j \|^2 \rightarrow \min \tag{2}$$

is represented as

$$\boldsymbol{R} = \boldsymbol{U} \cdot \mathrm{diag}(1, 1, \det\left[\boldsymbol{U}\boldsymbol{V}^\top\right])\boldsymbol{V}^\top \tag{3}$$

where the correlation matrix $\boldsymbol{K}$ is defined as $\boldsymbol{K} = \sum_j \boldsymbol{m}_j \boldsymbol{m'}_j^\top$ and decomposed as $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$ (where $\boldsymbol{U}$ and $\boldsymbol{V}$ are rectangular matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix).

Therefore, the 4×4 matrix $\boldsymbol{M}_{\mathrm{TW}}$, which maps the tag coordinate system to the world coordinate system, is expressed as $\boldsymbol{M}_{\mathrm{TW}} = \bar{\boldsymbol{T}_{\mathrm{W}}}\boldsymbol{R}\bar{\boldsymbol{T}_{\mathrm{T}}}^{-1}$ using the rotation matrix $\boldsymbol{R}$, the translation matrix $\bar{\boldsymbol{T}_{\mathrm{W}}}$ of $\bar{\boldsymbol{t}_{\mathrm{W}}}$ and the translation matrix $\bar{\boldsymbol{T}_{\mathrm{T}}}$ of $\bar{\boldsymbol{t}_{\mathrm{T}}}$. Then the pose $\boldsymbol{M}_{\mathrm{W}j}$ of the hidden background device in the world coordinate system is obtained as

$$\boldsymbol{M}_{\mathrm{W}j} = \boldsymbol{M}_{\mathrm{TW}}\boldsymbol{M}_{\mathrm{T}j} \tag{4}$$

D. The P4P method

In this subsection, we describe the second hidden background camera calibration method proposed in this study, which solves the Perspective 4-Points Problem using the relative pose between hidden background devices as a constraint, and uses the results to determine the poses of the hidden background devices (hereafter, the P4P method). In this subsection, we first describe the placement and measurement of the hidden background devices, and then describe the details of the P4P method.

In the P4P method, as in the intersection method, a situation is assumed where hidden background devices are placed behind obstacles at a certain distance from each other, but the number of hidden background devices to be placed is assumed to be greater than or equal to four, and at least four devices are placed on almost the same plane. While the intersection method requires a large user's area where the user can move around with the measuring devices, the P4P method can be used even when the user's area is small.

In the following, we explain the case where four hidden background devices are placed as an example. As the Step 1 in the intersection method, after placing the hidden background devices, a checkerboard is used to determine the relative pose between the hidden background devices.

Next, as the Step 2, as in the intersection method, the direction of placement of the hidden background devices is measured using the measuring device.

Then, as the Step 3, the pose of the measuring device is obtained in the same way as for the intersection method.

Then, as in the intersection method, the Step 4 is to divide the space where the measuring device was moved during the measurement into equally spaced grids, and accumulate the directional information for each grid.

Finally, as the Step 5, the poses of the hidden background devices in the world coordinate system is estimated by solving the Perspective-n-Point problem (PnP problem). The process flow is shown in Fig. 4.

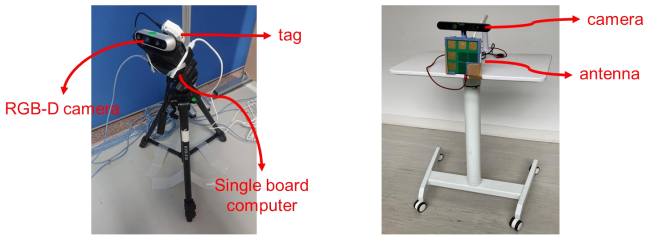
In the P4P method, the P4P problem is first solved for each grid to obtain the pose $\boldsymbol{M}_{\mathrm{W}ij}$ of the hidden background device j in the world coordinate system. However, since the P4P problem targets 3D points captured by a camera, and what is obtained by the measuring device used in this study

is the direction vectors of the hidden background devices, the solution method of the P4P problem cannot be used as it is. Therefore, in this study, a virtual camera is set up, the directions of the hidden background devices are projected onto the image of the virtual camera, and the result is used to solve the P4P problem. First, assume a virtual camera whose focal point is the center coordinate of the grid and whose shooting direction is the average of the directions in which the four hidden background devices exist. The focal length of the virtual camera and the resolution of the image should be set appropriately so that all the hidden background devices are visible within the shooting range. Next, calculate the two-dimensional coordinates of the hidden background devices as seen from the virtual camera when projected onto the image plane. The pose $\boldsymbol{M}_{\mathrm{W}ij}$ of the hidden background device j in the world coordinate system is obtained by solving the P4P problem from these two-dimensional coordinates and the previously measured relative positions of the cameras. The above operations are performed for all grids. Next, the following operations are performed to select the one that is considered to be the most accurate from $\boldsymbol{M}_{\mathrm{W}ij}$ obtained for each grid. First, the 3D position $\boldsymbol{W}_{ij}$ in the world coordinate system of the hidden background device j is obtained using the measurement results in grid i. Next, use the measurement results on the other grid i' to find the 3D position $\boldsymbol{W}_{i'j}$ in the world coordinate system of the hidden background device j, and find the Euclidean distance between $\boldsymbol{W}_{ij}$ and $\boldsymbol{W}_{i'j}$. After that, find the sum of the number of times n_i the measuring device passed through the other grid i' during the measurement, whose value is less than a predetermined threshold for all hidden background device $(j = 0, 1, 2, 3)$. The value of n_i is considered to represent how reliable the measurement results of grid i are when judged based on the measurement results in other grids. Therefore, the 3D pose of the hidden background device obtained from the grid with the highest n_i value is the final result of the hidden background device obtained by the P4P method.

IV. Experiments

A. Overview

Experiments were conducted to evaluate the accuracy of the hidden background device pose estimated using the intersection method and the P4P method. In this evaluation, the true values of the poses of the hidden background devices were first measured manually. Next, we estimated poses of the hidden background cameras using the intersection method and the P4P method, and determined the error. As measures of the error, we used the difference $\Delta x, \Delta y, \Delta z$ between the true value of the hidden background device position and the difference $\Delta\beta, \Delta\alpha, \Delta\gamma$ between the true value and the estimated value when the orientation of the hidden background device was expressed in terms of the y-x-z system Euler angles. Finally, we compared the synthesized images for the DR experience using the poses of the hidden background devices estimated by the intersection method and the P4P method with the results of synthesized images using the true values.



The hidden background device The measuring device

Fig. 5. Picture of the experimental devices

B. Hardware configurations

Fig. 5 shows the hardware used for the evaluation: the hidden background device and the measuring device.

The hidden background device consists of a tag for direction finding, an RGB-D camera, a battery and a single board computer to drive the camera. In this evaluation, the Direction Finding function of Bluetooth 5.1 was used as the wireless direction finding technology, and the C209 tag of u-blox' XPLR-AOA-1 kit [10] was used as the tag. The angle accuracy of XPLR-AOA-1 is 5° mean error and the update rate is 40 updates per second. For the hidden background camera, we used Intel RealSense D435. A Raspberry Pi 4 Model B from the Raspberry Pi Foundation was used as the single board computer.

Finally, we describe the measuring device that simultaneously photographs the user's area and detects the directions of radio waves from the hidden background devices. The measuring device consists of a camera that takes pictures of the user's area and an antenna that detects the directions of the tags. We used ASUSTek Computer's Xtion Pro Live as the camera to take pictures of the user's area. A C211 anchor from the XPLR-AOA-1 kit was used as the antenna to detect the directions from the radio sources.

C. Experimental environment

A floor plan of the room used for the evaluation is shown in Fig. 6. The user's area and the hidden background area were separated by fabric partitions.

Next, we describe the placement conditions of the hidden background devices. The number of the hidden background devices was set to four, and as shown in Fig. 6, all devices were placed on approximately the same plane with all devices facing in the direction of shooting the far side of the hidden background area. When capturing the user's area with the measuring devices, the measuring devices were moved within the area indicated by rectangle ABCD in the figure. From this region, the hidden background devices were not directly visible and the directions of the devices could only be detected using the antenna.

D. Measuring in user's area

The poses of the hidden background devices were measured using the measuring device. In this measurement, the RGB-D camera of the measuring device is used to capture images of the user's area, and the antenna is used

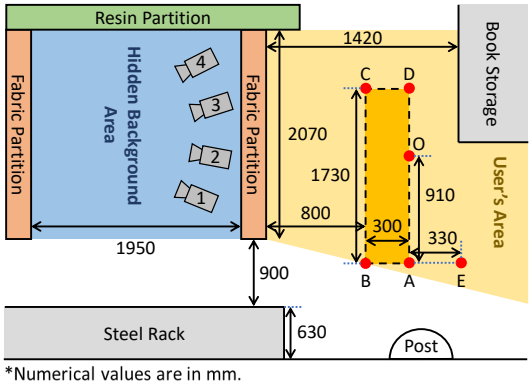


Fig. 6. Layout of the experimental room

to measure the directions of radio waves from the hidden background devices. The images obtained using the RGB-D camera are used to perform SfM, and the pose of the RGB-D camera when each image is taken is determined, and the result is used as the pose of the antenna at that time. In this measurement, the measuring device was moved in the following three ways, and the measurements and images were taken for 1800 frames.

- The measuring device is moved within the rectangular ABCD area shown in Fig. 6, varying in height (approximately 0.7 m to 1.1 m from the floor) and direction (free movement).
- The measuring device is moved from point A to point D, as shown in Fig. 6, without changing direction or height (linear movement).
- The measuring device fixed at a height of about 0.85 m from the floor at the point O shown in Fig. 6 (static).

The intersection method and P4P method were applied to the results measured by each movement to obtain the poses of the hidden background devices.

E. A method for creating DR experience images

The object to be removed using DR was the central fabric partition shown in Fig. 6, and the visualization target is a person standing in the hidden background area. The partition was taken from the location of point E in Fig. 6 (about 0.85 m high from the floor) using the same RGB-D camera as the measuring device, and SfM was performed on the resulting images to obtain the pose of the RGB-D camera. The person standing in the hidden background area was photographed using hidden background devices, and a colored point cloud was generated with reference to each hidden background device using the color and depth images obtained. Then, using the pose of the RGB-D camera obtained by SfM and the poses of the hidden background devices obtained by using the intersection method and the P4P method, the colored point cloud was projected as the color image. At that time, a rectangular area corresponding to a virtual window to look into the back side was set on the partition beforehand, and the inside of the rectangle was first filled with black, and then only the colored point cloud projected in the rectangle was superimposed on the final image.

F. Implementation of the proposed methods

All application were implemented in C++ language using Microsoft's Visual Studio as a development environment. As a library, we used Point Cloud Library (Ver. 1.12.0) [13] for the point cloud display of the DR experience. The grid size was set to 50 mm. The focal length of the virtual camera in the P4P method was 500 pixels on both the horizontal and vertical axes, and the resolution of the image was 1000×1000 pixels.

V. Results and Discussion

First, we describe the evaluation of numerical errors in the positions and orientation angles of the hidden background devices. Table I shows the difference between the true values and the resultant poses of the hidden background devices obtained using the intersection method and the P4P method under free movement, linear movement and static conditions. In the free movement condition, the P4P method was able to estimate the pose of the hidden background device relatively accurately with positional errors of 100 mm or less in all xyz axes and orientational errors of 10 degrees or less in all axes, whereas the intersection method had larger errors in both. The reason for this large error in the intersection method is thought to be that the distance between the points AD where the measuring devices were moved during the measurement and the distance between each hidden background device were too short, which did not reduce the error in the radio detection. This suggests that one way to improve the accuracy of the pose of the hidden background device by the intersection method is to increase the distance between each hidden background device so that each intersection appears more clearly with respect to the others.

In the linear movement condition, the position and orientation errors of the P4P method were large, while the errors of the intersection method were very small compared with the free movement condition. One possible reason for the smaller error in the intersection method is that the number of measurement frames was standardized at 1,800 frames for all three measurements this time, and compared to the free movement condition, the elimination of vertical movement of the measuring device allowed it to pass through a more limited area, resulting in more data measurements per grid and a reduction in tag direction error. Therefore, when performing calibration using the intersection method, if the number of data measurements needs to be reduced, the poses of the hidden background devices can be obtained with a high degree of accuracy by moving the measuring devices significantly in only one direction.

In the static condition, the intersection method does not work because the measuring device does not move at all. On the other hand, the P4P method can determine the poses of the hidden background devices even if the measuring device does not move at all during measurement and only one frame of data is obtained. However, the error was larger than in the free movement and linear movement conditions, in which data is measured while the measuring device is moved. In order to perform accurate calibration using the P4P method

TABLE I

ERROR BETWEEN MEASURED HIDDEN BACKGROUND DEVICE POSES AND TRUE VALUE (INT: INTERSECTION METHOD, P4P: P4P METHOD)

	Free movement		Linear movement		Static	
	Int	P4P	Int	P4P	Int	P4P
Δx(mm)	8.6	-61.2	19.0	-40.9	3675.0	-41.8
Δy(mm)	-356.4	-7.0	-164.5	-50.9	286.3	-140.9
Δz(mm)	-641.0	-99.1	-291.5	-192.2	2789.1	-897.6
$\Delta\beta(^\circ)$	69.5	-1.0	-16.3	-12.5	53.3	-29.1
$\Delta\alpha(^\circ)$	40.3	3.6	12.7	3.9	2.3	-14.4
$\Delta\gamma(^\circ)$	98.1	-9.3	8.3	-6.8	158.0	-15.4

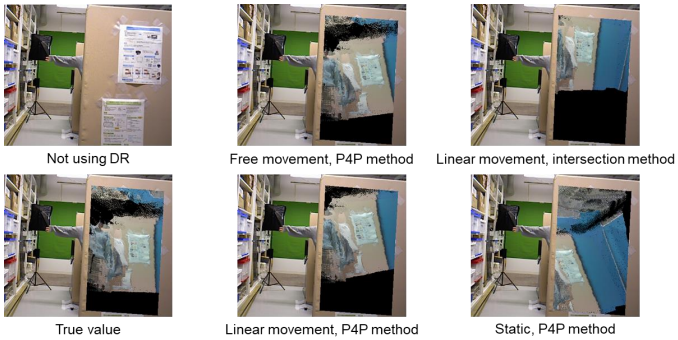


Fig. 7. Example of execution of DR (using true and measured value of the pose of the hidden background device)

with fixed measuring devices, it is necessary to increase the distance between the hidden background devices as much as possible, and increase the time to take pictures with the device. In addition, instead of accumulating the data for each grid, the poses of the hidden background devices in the world coordinate system should be calculated for each frame of data, and the best poses among them should be chosen from the results.

Next, we describe the results of generating DR experience images using the poses of the hidden background devices obtained by each moving and calibration method. A DR experience image created using the true value is shown on the left side of Fig. 7. Also, the DR experience image created using the poses of the hidden background devices obtained by each moving method is shown on the right side of Fig. 7. Comparing the images using true value with measured value using the P4P method under the free movement condition in Fig. 7, the error in the position of the hidden background device is not very noticeable on the images, whereas the error in the rotation around the z-axis is highly noticeable. In addition, even when the error of rotation around the y-axis is large, as in the images using the P4P method under the static condition, the hidden background area is not correctly visualized. These indicate that errors in orientation, especially errors in rotation around the z-axis, are more noticeable than errors in estimating the positions of the hidden background devices. Therefore, it is considered to be a future issue to reduce the error in the rotation around the z-axis of the shooting direction.

VI. CONCLUSION

To develop a method for calibrating hidden background cameras through obstacles, we proposed an idea in which a hidden background camera is equipped with a wireless radio source and calibrated by detecting the direction of the radio waves. As a method to obtain the pose of the hidden background camera from the directional detection of radio waves containing errors, we proposed the intersection method, which determine the positions of the hidden background devices from the intersection of vectors in the direction in which the devices are located, and the P4P method, which uses the relative pose of hidden background devices as a constraint to determine the poses of the hidden background devices. The results of the method evaluation show that the P4P method provides better accuracy. It was also found that while errors in the positions of the obtained hidden background devices have a relatively small impact on the DR experience image, errors in the orientation have a very large impact. In particular, the error of rotation around the optical axis of the camera, even if small, is likely to cause a sense of discomfort in the DR experience.

Future work includes improving the technique to suppress the error of rotation around the optical axis of the camera and to further improve the robustness against outliers.

REFERENCES

[1] S. Mori, S. Ikeda and H. Saito: A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects, IPSJ Trans. on Computer Vision and Applications, 2017.

[2] K. W. Cheung and H. C So: A multidimensional scaling framework for mobile location using time-of-arrival measurements, Signal Process, vol. 53, pp. 460-470, 2005.

[3] J. Xu, M. Ma and C. L. Law: Position Estimation Using UWB TDOA Measurements, 2006 IEEE Int. Conf. Ultra-Wideband, Waltham, MA, USA, pp. 605-610.

[4] K. Chawla, C. McFarland, G. Robins and C. Shope: Real-time rfid localization using rss, Int. Conf. Localization and GNSS, pp.1–6, 2013.

[5] Q. Fengzhong, W. Shiyuan, W. Zhihui and L. Zubin: A survey of rangingalgorithms and localization schemes in underwater acoustic sensor network, China Commun., vol. 13, no. 3, pp. 66–81, 2016.

[6] T. Boroushaki, M. Lam, L. Dodds, A. Eid and F. Adib: Augmenting Augmented Reality with Non-Line-ofSight Perception, 20th USENIX Symp. Networked Systems Design and Implementation, pp. 1341-1358, 2023.

[7] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas and M.J. Marín-Jiménez: Automatic generation and detection of highly reliable fiducial markers under occlusion, Pattern Recognition, Vol.47, Issue.6, pp.2280-2292, 2014.

[8] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas and R. Medina-Carnicer: Generation of fiducial marker dictionaries using Mixed Integer Linear Programming, Pattern Recognition, Vol.51, pp.481-491, 2016.

[9] G. Wahba: A Least Squares Estimate of Satellite Attitude, SIAM Review, Vol.7, No.3, pp.409, 1965.

[10] XPLR-AOA-1 kit （2023, January 29). https://www.u-blox.com/en/

[11] Detection of ChArUco Boards (2023, February 4). https://docs.opencv.org/3.4/df/d4a/tutorial_charuco_detection.html/

[12] X. Zhi, J. Hou, Y. Lu, L. Kneip and S. Schwertfeger: Multical: Spatiotemporal Calibration for Multiple IMUs, Cameras and LiDARs, 2022 IEEE/RSJ Int. Conf. Intelligent Robots and Systems, Kyoto, Japan, pp. 2446-2453.

[13] R. B. Rusu and S. Cousins: 3D is here: Point Cloud Library (PCL), 2011 IEEE Int. Conf. Robotics and Automation, Shanghai, China, pp. 1-4.