A Method to Estimate User Position and Orientation Based on Relative Position Measurement Between a Navigation Robot and a Smartphone

Satoki Ogiso
National Institute of Advanced Industrial Science and Technology (AIST)
Kashiwa, Japan
Email: s.ogiso@aist.go.jp

Daisuke Sato
Carnegie Mellon University
Pittsburgh, United States
Email: daisukes@cmu.edu

Ryosuke Ichikari, Takahiro Miura, Masakatsu Kourogi, Takashi Okuma, Takeshi Kurata, Takeshi Tanabe
National Institute of Advanced Industrial Science and Technology (AIST)
Kashiwa, Japan
Email: {r.ichikari, miura-t, m.kourogi, takashi-okuma, t.kurata, t-tanabe}@aist.go.jp

***Abstract*—This paper proposes a method to estimate a smartphone's location and orientation using the relative position between it and a navigation robot. The relative position is measured by ultra-wideband (UWB) ranging and angle of arrival (AoA). The smartphone and navigation robot estimate their positions in their coordinate systems. The proposed method estimates the transformation of these two coordinate systems using UWB measurements. The localization error was evaluated by an experiment in a museum with six participants. As a result, the localization error of the proposed method was 0.40 m in the median, which is better than 0.65 m, the error with a baseline method that aligns coordinates at the start of localization. This method enables higher accuracy of smartphone localization for the users of navigation robots in short-term leave.**

***Index Terms*—Ultra-wideband, Visual-inertial odometry, Navigation robot, Relative position, Extended Kalman filter**

I. Introduction

Support for orientation and mobility is crucial for assisting individuals with visual impairments. Navigation assistants with mobile devices or smartphones have been studied [1], [2]. In this context, various methods for position estimation have been introduced, including those based on pedestrian dead reckoning (PDR) [3], Bluetooth low energy (BLE) [4], QR codes [5], AR markers with visual-inertial odometry and their combinaitons [6], [7]. The PDR and BLE-based approach does not require initialization but has a relatively low position error of approximately 1 meter. Methods using QR codes or AR markers require the user to point the camera toward the marker for initialization, which may be challenging to complete for people with visual impairments. Also, other ultra-wideband (UWB) [8], [9] or 3D map-based matching methods [10] require the setup and maintenance of the infrastructures.

Recent research enabled navigation robots for guidance [11], [12]. This approach offers relatively high-accuracy position estimation as long as the user walks with the robot. The user, however, may want to leave away from the robot to navigate in their destination space alone, which may necessitate additional navigation methods with small mobile devices and higher accuracy to explore the environment (i.e., finding items, appreciating exhibition content.)

One of the solutions for this problem could be cooperative localization among multiple localization agents. There are several types of research on cooperative localization of multiple autonomous vehicles [13], [14]. To the authors' best knowledge, cooperative localization between a robot and a smartphone for pedestrian navigation has not been investigated.

This paper is based on results obtained from a project, Programs for Bridging the gap between R&D and the IDeal society (society 5.0) and Generating Economic and social value (BRIDGE)/Practical Global Research in the AI × Robotics Services, implemented by the Cabinet Office, Government of Japan.

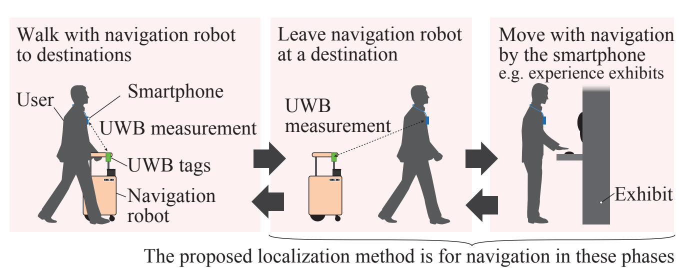

Fig. 1. A usage scenario of the proposed method.

Therefore, we propose a method to estimate the smartphone's location by utilizing a navigation robot as a reference point using relative position measurement by UWB. Fig. 1 shows the usage scenario of the proposed method. Since a navigation robot can contain various sensors and estimate its location with relatively higher accuracy, the proposed system utilizes its location as a reference point. Multiple UWB tags are attached to the robot, and the smartphone and UWB tags measure their UWB signal to each other to estimate their relative coordinate. It allows for position estimation without pointing the smartphone in a specific direction or adding major infrastructural modifications to the facility.

II. Proposed localization method

A. Overview

The proposed method assumes that the smartphone and the navigation robot have their localization systems and coordinate systems O^A and O^R. In this paper, the smartphone does not have a map of the environment and runs visual-inertial odometry (VIO). This assumption is made because the smartphone can be hung from the user's neck during navigation, and the camera is available. While VIO is accurate in a small region, it is susceptible to drift and does not know the global position without other information. The distance and the relative angle between the phone and the navigation robot are measured using ultra-wideband (UWB). Fig. 2 shows the overview of the proposed method, and Fig. 3 shows an illustration of the variables. Definitions and descriptions of the variables are in Table I. The navigation robot has a map and is equipped with sensors such as LiDAR to measure its position and orientation. The UWB tags are equipped on the navigation robot. The position and orientation of the UWB tags are calculated by the navigation robot's localization.

The proposed method uses an extended Kalman filter (EKF) to estimate the state vector $\boldsymbol{z}_A^R = [\boldsymbol{x}_A^R, \theta_A^R]^{\boldsymbol{T}}$ with the position $\boldsymbol{x}_A^R$ and yaw angle θ_A^R. Also, the covariance matrix corresponding to this state variable is P_A^R. In the prediction step, the origin of the coordinate system is assumed not to move, and predictions are made so that only the covariance increases with time as follows.

$$P_A^R \leftarrow P_A^R + Q \tag{1}$$

where Q is the covariance matrix, assuming a normally distributed noise is added to the position-posture. The proposed method assumes that the gravity vector aligns the vertical direction of the frames A and R. Thus, only the yaw angle represents the transformation between the navigation robot and the phone's coordinates.

B. EKF update with UWB measurement

Assume that the UWB tags and the phone can estimate the distance between them and the angle of arrival (AoA) at the receiver. The phone-side measurement can be expressed as the following observation function:

$$\boldsymbol{d}_{U_i}^P = T\left((\boldsymbol{q}_{U_i}^R)^{-1}\right)\left(\boldsymbol{x}_A^R + T\left(\boldsymbol{q}_A^R\right)\boldsymbol{x}_P^A - \boldsymbol{x}_{U_i}^R\right) \tag{2}$$

$$\boldsymbol{h}_{U_i}^P(\boldsymbol{z}_A^R) = \left[|\boldsymbol{d}_{U_i}^P|, \boldsymbol{d}_{U_i}^P / |\boldsymbol{d}_{U_i}^P|\right] \tag{3}$$

TABLE I

VARIABLES AND THEIR DEFINITIONS.

Variable	Definition / Description
O^A	Origin of the phone's local coordinate system, referred to as phone frame A
O^R	Origin of the navigation robot's coordinate system, referred to as reference frame R
$\boldsymbol{x}_A^R$	Position of O^A in the frame R.
θ_A^R	Yaw angle of O^A in the frame R.
$\boldsymbol{z}_A^R$	State vector of the EKF composed of the $\boldsymbol{x}_A^R$ and θ_A^R
$\boldsymbol{q}_A^R$	Orientation (quaternion) of O^A in R. In this paper, it is yaw angle θ_A^R expressed in the quaternion form.
P_A^R	Covariance matrix of the state estimate $\boldsymbol{z}_A^R$.
$\boldsymbol{q}_P^A$	Orientation of the phone in the frame A.
$\boldsymbol{x}_P^A$	Position of the phone the frame A.
$\boldsymbol{x}_{\mathrm{GT}}$	Ground truth of the phone's position in the frame R.
$\boldsymbol{x}_{U_i}^R$	Position of the ith UWB tag in the reference frame R.
$\boldsymbol{q}_{U_i}^R$	Orientation of the ith UWB tag in the reference frame R.
K	Kalman gain
$R_P^{U_i}$	Measurement noise covariance of tag side measurement at ith tag.
$R_{U_i}^P$	Measurement noise covariance of phone side measurement of ith tag.
Q	Process noise covariance of EKF.
$\boldsymbol{h}_{U_i}^P(\boldsymbol{z}_A^R)$	Observation function that maps the state vector to the range and directional unit vector which points ith UWB tag in the phone-fixed frame P.
$\boldsymbol{h}_P^{U_i}(\boldsymbol{z}_A^R)$	Observation function that maps the state vector to the range and directional unit vector which points the phone in the tag-fixed frame U_i.
$H_{U_i}^P(\boldsymbol{z}_A^R)$	Jacobian of $\boldsymbol{h}_{U_i}^P(\boldsymbol{z}_A^R)$.
$H_P^{U_i}(\boldsymbol{z}_A^R)$	Jacobian of $\boldsymbol{h}_P^{U_i}(\boldsymbol{z}_A^R)$.
$\boldsymbol{d}_{U_i}^P$	Position of the ith UWB tag expressed in the frame P.
$\boldsymbol{d}_P^{U_i}$	Position of the phone expressed in the frame U_i.
$\boldsymbol{y}_{U_i}^P$	Measurement of UWB range and AoA that corresponds to $\boldsymbol{h}_{U_i}^P(\boldsymbol{z}_A^R)$
$\boldsymbol{y}_P^{U_i}$	Measurement of UWB range and AoA that corresponds to $\boldsymbol{h}_P^{U_i}(\boldsymbol{z}_A^R)$
$T(\boldsymbol{q})$	Transformation (rotation) matrix associated with the quaternion $\boldsymbol{q}$.
T_d	Delay of the UWB siganl evaluation.

This equation corresponds to the illustration in Fig. 3(b). Since the principle differs for distance and AoA measurements, we would like to set the covariance by breaking it down into distance and AoA, rather than covariance along the coordinate axis of relative position. Thus, the Eq. (2) was transformed into Eq. (3) to separate ranging and AoA. The update step is based on each of the following observation equations. Let $\boldsymbol{H}_{U_i}^P(\boldsymbol{z}_A^R)$ be the Jacobian of Eq. (3). Also, let $R_{U_i}^P$ be the covariance row example with the variance of the ranging and the variance of the AoA angle as diagonal components. The state variables

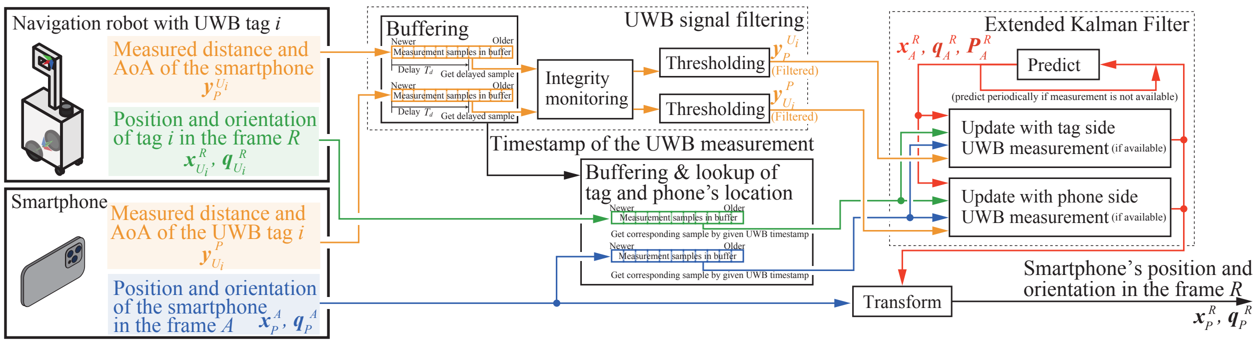

Fig. 2. Overview of the proposed method. The proposed method estimates the transformation of the smartphone's position and orientation from its frame A to the navigation robot's frame R. UWB signals are filtered in advance and evaluated with a delay to wait until the corresponding navigation robot's position is available.

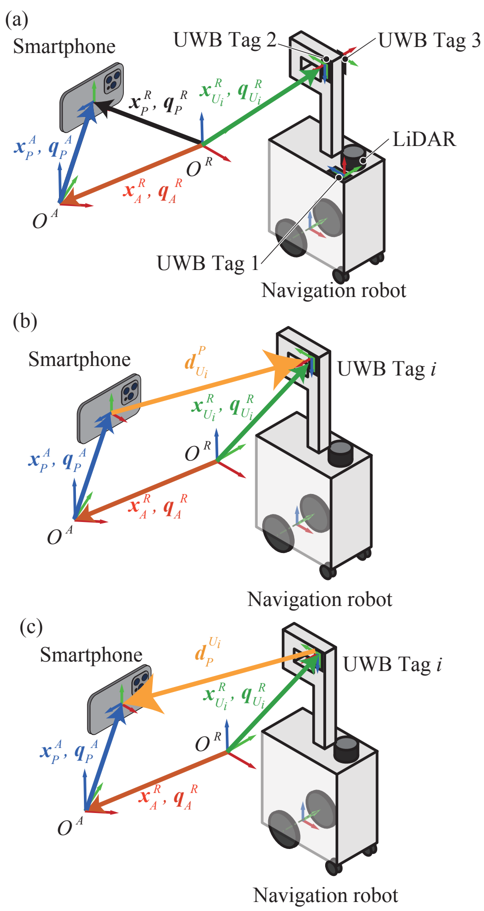

Fig. 3. Description of the variables. (a) Definition of the coordinate origins and the locations, (b) an illustration of the UWB measurement on the tag side, and (c) that of on the phone side. This paper aims to estimate $\boldsymbol{x}_A^R$ and $\boldsymbol{q}_A^R$ (red transformation in (a)) to obtain the position of the phone in the reference frame R.

and estimates of their variances are then updated as follows.

$$K = \frac{P_A^P H_{U_i}^P(\boldsymbol{z}_A^R)^{\mathrm{T}}}{H_{U_i}^P(\boldsymbol{z}_A^R) P_A^P H_{U_i}^P(\boldsymbol{z}_A^R)^{\mathrm{T}} + R_{U_i}^P} \quad (4)$$

$$\boldsymbol{z}_A^R \leftarrow \boldsymbol{z}_A^R + K(\boldsymbol{y}_{U_i}^P - \boldsymbol{h}_{U_i}^P(\mathrm{z}_\mathrm{A}^\mathrm{R})) \quad (5)$$

$$P_A^R \leftarrow (I - K H_{U_i}^P(\boldsymbol{z}_A^R)) P_A^R \quad (6)$$

$\boldsymbol{y}_{U_i}^P$ is the measurement of UWB range and AoA that corresponds to Eq. (3).

Similarly, the tag side measurement can be expressed as the following observation function. This equation corresponds to the Fig. 3(c).

$$\boldsymbol{d}_P^{U_i} = T\left((\boldsymbol{q}_{u,i}^R)^{-1}\right)\left(\boldsymbol{x}_A^R + T\left(\boldsymbol{q}_A^R\right)\boldsymbol{x}_p^A - \boldsymbol{x}_{u,i}^R\right) \quad (7)$$

$$\boldsymbol{h}_P^{U_i}(\boldsymbol{z}_A^R) = \left[|\boldsymbol{d}_P^{U_i}|, \boldsymbol{d}_P^{U_i}/|\boldsymbol{d}_P^{U_i}|\right] \quad (8)$$

Let $\boldsymbol{H}_P^{U_i}(\boldsymbol{z}_A^R)$ be the Jacobian of Eq. (8), $R_P^{U_i}$ be the covariance and $\boldsymbol{y}_P^{U_i}$ be the measurement corresponding to Eq. (8). The update step is expressed as follows.

$$K = \frac{P_A^P H_P^{U_i}(\boldsymbol{z}_A^R)^{\mathrm{T}}}{H_P^{U_i}(\boldsymbol{z}_A^R) P_A^P H_P^{U_i}(\boldsymbol{z}_A^R)^{\mathrm{T}} + R_P^{U_i}} \quad (9)$$

$$\boldsymbol{z}_A^R \leftarrow \boldsymbol{z}_A^R + K(\boldsymbol{y}_P^{U_i} - \boldsymbol{h}_P^{U_i}(\mathrm{z}_\mathrm{A}^\mathrm{R})) \quad (10)$$

$$P_A^R \leftarrow (I - K H_P^{U_i}(\boldsymbol{z}_A^R)) P_A^R \quad (11)$$

UWB measurements are once stored in a buffer and later evaluated. The buffer was periodically checked, and if a UWB value was observed, the above update process was performed.

C. UWB signal filtering

1) Integrity monitoring between tag-side and phone-side measurement: Especially when using a smartphone, it may not be possible to access the raw UWB measurement result. For example, in iOS 18.2, the UWB ranging sometimes shows as if the tag is stopped even though it is moving with the phone (see Fig. 7 for example). The proposed method verifies that the information on the smartphone and the tag sides are appropriately matched. The pairs of rangings by the tag and smartphone are checked, and if the difference in the ranging results is above a threshold, the proposed method discards the ranging sample. Also, the tag has a built-in non-line-of-sight (NLOS) estimation. The proposed method ignores the tag side measurements with NLOS class and corresponding phone-side measurements. In this paper, the threshold was set at 0.2 m.

2) Range and AoA threshold: The farther away from the tag, the more susceptible it is to multipath, and the AoA error becomes magnified when converted to location. Thus, the proposed method does not run EKF update if the phone and navigation robot are too far. Also, due to the characteristics of the antenna, the resolution of AoA is poor when the angle is near the straight line created by the antenna array. The proposed method limits valid AoA angles. The measurements from the UWB tags are obtained for azimuth and elevation, each of which is calculated by a pair of antennas. Since these measurements are from independent antenna arrays, we set an upper limit for each, and only when they are below a threshold value, they are converted to direction vectors and used for updating the EKF. On the other hand, it is unclear whether the above decomposition into azimuth and elevation is effective for the smartphone side since the UWB antenna placement has not been disclosed. Therefore, the angle that the obtained direction vector makes with respect to the frontal direction vector was used as a criterion. Only data for which this does not exceed a threshold value were used to update the EKF. In this paper, the range threshold was 5 m. The threshold for azimuth and elevation for tag side AoA and the direction vector angle for smartphone side AoA was ± 50 degrees.

3) Delayed evaluation: This research used the UWB tag measurement for EKF update with a certain delay for two reasons. Firstly, UWB tag position and orientation are sequentially obtained from the robot's self-localization result. However, since estimating the robot's position takes some time, the UWB tag positions cannot be used as soon as the UWB measurement is obtained. The typical delay of the robot's position estimation is around 0.1 to 0.4 seconds in the experiment. For this reason, the received information is not used immediately but is updated in EKF after a certain period of time. Since the state variable of EKF is the origin on the smartphone side, position and orientation estimation of the phone can be continued in real time without problem even if the update is delayed. Secondly, the UWB measurement becomes inaccurate when transitioning from line-of-sight to non-line-of-signt. To ignore the transitioning data, the proposed method only uses UWB measurement, which continuously has measurements at an expected rate for a period before and after that sample. For evaluating continuity after the sample, the proposed method evaluates the UWB measurement with a certain delay. In this paper, the delay was set at 2.0 seconds to ensure UWB sample is not measured in the LOS/NLOS transition condition.

III. Experiment

A. Equipment for the proposed method

An AI Suitcase, which was made specifically for this experiment, was used as the navigation robot. For position and orientation estimation, the navigation robot was equipped with a LiDAR sensor (XT-16, Hesai), an inertial measurement unit (BNO055, Bosch), and a barometer (BME280, Bosch). The experimental environment was mapped in advance using the Cartographer[1], and BLE reception conditions were investigated. During position estimation, BLE reception was used to identify the floor, and the robot's position and orientation on that floor's map were estimated. The navigation robot was equipped with UWB tags (firmware version v04.06.00, Type 2BP Evaluation Kit, Murata Manufacturing) with a sampling frequency of around 5 Hz. Three UWB tags were placed on the navigation robot as shown in Fig. 2(a). An overview of the equipment is shown in Fig. 4. The NLOS estimation method is the proprietary method by the manufacturer with the firmware version. A smartphone (iOS18.2, iPhone 13 Pro, Apple) was used for visual SLAM via ARKit without needing a map. Distance and angle information were obtained via the iOS Nearby Interaction API without the "camera assistance" option. Also, tag-side UWB measurements by the UWB tags are recorded on the navigation robot. The proposed method will be evaluated by comparing the estimated position of the phone and the ground truth obtained by optical-based localization. All measured data in the phone was transferred to the navigation robot via MQTT and stored with other data. The results in this paper are processed after the experiment with constraints as if it is a real-time execution. Initial value of P_A^R was $\mathrm{diag}[10, 10, 10, 100]$. The covariances are set as $R_P^{U_i} = \mathrm{diag}[1, 10, 10, 10]$, $R_{U_i}^P = \mathrm{diag}[1, 10, 10, 10]$, $Q = 10^{-5} I_{7\times 7}$.

[1]https://github.com/cartographer-project

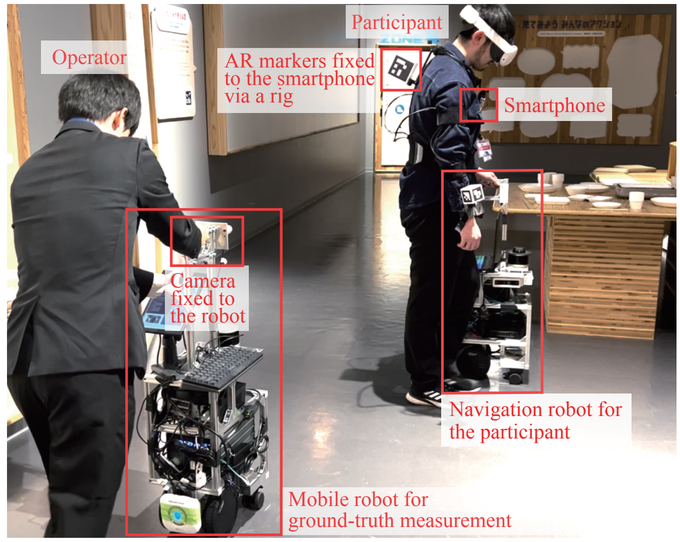

Fig. 4. Example of the experimental situation. The participant wears a smartphone and AR markers and is guided by a navigation robot. The operator follows the participant with the camera-equipped mobile robot to capture AR markers for the ground truth of the smartphone location.

B. Experiment location and paths

The experiment was conducted in a museum (Miraikan, The National Museum of Emerging Science and Innovation, Tokyo, Japan) while the exhibits were open to the public. An overview of the exhibit floor is shown in Fig. 5(a). Each exhibit starts at the circle with the number "1," stops at the points marked with numbers in circles, and then heads to the exit of each exhibit, viewing the exhibits as they go. Exhibit 1 shown in Fig. 5(b) has structures mostly made of wood and no significant blockage of the UWB. Exhibit 2 shown in Fig. 5(b) has small rooms with curtains, which causes

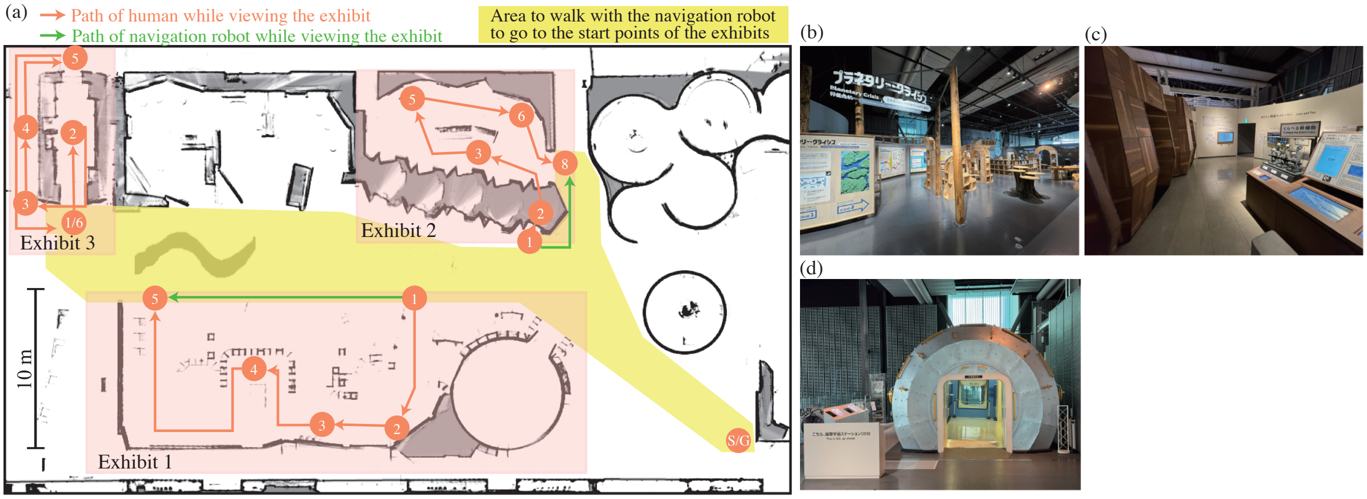

Fig. 5. Map of the experimental environment. (a) Top view of the museum floor and paths inside exhibits. Note that the participant starts from the right-bottom start/goal position and walks with the navigation robot to each exhibit. (b) Exhibit 1 is mostly open space with several wooden structures. (c) Exhibit 2 has small rooms with curtains at the start of the exhibit. (d) Exhibit 3 is a mock-up of the International Space Station made of metal.

occlusion to the smartphone's visual-inertial system. Exhibit 3 shown in Fig. 5(b) is a real-scale International Space Station module made mostly of metal, which causes severe blockage of UWB signals. The participants went through predetermined routes for each exhibit shown in Fig. 5(a). The navigation robot guided them to the start of each exhibit route. They were asked to walk under the guidance of the navigation robot or walk apart from the robot by following the voice guidance of the experiment staff, who mimicked the voice navigation of a smartphone. The participants touch the exhibits with their hands, etc., assuming that they were guided through the exhibits. Each participant wore a phone at the neck and walked in the environment with the navigation robot. The participants are four sighted and two participants with visual impairments. The total duration of the data processed in this paper is 139 minutes. The experiment was approved by the IRB committee of AIST (AIST Ergonomic Experiment Committee), the first author's organization (HF2024-1438).

C. Ground truth and a baseline method for evaluation

To evaluate the error of position estimation using the proposed method, ground truth measurement was performed using AR markers. Fig. 4 shows the overview of this process. The smartphone and AR markers were attached to a rig that could be regarded as a rigid body. The smartphone was placed on the subject's chest, and a cube with AR markers on five sides was placed on the subject's back. An operator held a mobile robot dedicated to ground truth measurement from behind and moved it so that the AR markers appeared in the field of view of the camera (D455, Intel). The position and orientation of the AR markers, as seen from the camera, and the positional posture of the camera obtained by the ground truth measurement robot were converted to the position and orientation of the smartphone. The size of the AR marker is 8×8 cm and used the dictionary of DICT_5X5_50 of ArUco markers [15]. The mobile robot for the ground truth has the same LiDAR, IMU, and barometer as the AI suitcase for the participant. The localization samples are regarded as outliers and removed if the spacing of the sample before and after are above 1 second, or velocities are above 2 m/s.

The correctness of ground truth by AR markers was evaluated a priori. The rig, which was integrated with a smartphone and AR marker, was aligned to the corners of the map to be located, and the deviations of the positions from the corners were compared. In this paper, the localization error is evaluated with circular error (CE). It is the error in the horizontal plane, which is an important criterion of applications for pedestrians [16], [17]. CE50 stands for the 50 percentile of the circular error and CE 90 for the 90 percentile. CE50 of the AR marker-based ground truth is 0.087 m in the horizontal plane against the map's corners. Based on this result, we decided to use it as the ground truth, assuming it is sufficiently correct for the assumed positioning method.

The proposed method was compared with a baseline method that imitates the alignment of the position and orientation at the start of the system. The baseline method uses the 30 seconds of the walk and estimates the best-fitting transformation by solving the following optimization.

$$\boldsymbol{x}_A^R, \theta_A^R = \underset{\boldsymbol{x}_A^R,\theta_A^R}{\operatorname{argmin}} \sum_{T_s<t<T_e} \left|\boldsymbol{x}_{\mathrm{GT}}(t) - T(\boldsymbol{q}_A^R(\theta_A^R))\boldsymbol{x}_P^A(t) + \boldsymbol{x}_A^R\right|^2 \tag{12}$$

Where $\boldsymbol{x}_{\mathrm{GT}}(t)$ is the ground truth location at time t in the frame R. Equation (12) was solved by L-BFGS (Limited-memory Broyden–Fletcher–Goldfarb–Shanno) method. The localization result of the baseline method is obtained by transforming the phone's local location $\boldsymbol{x}_A^R$ with this best-fitting transformation $\boldsymbol{x}_A^R, \theta_A^R$. Note that this baseline uses ground truth values, which might be replaced in other methods with AR markers at a fixed point, feature-based visual localization, or fixing the phone at a pre-determined location.

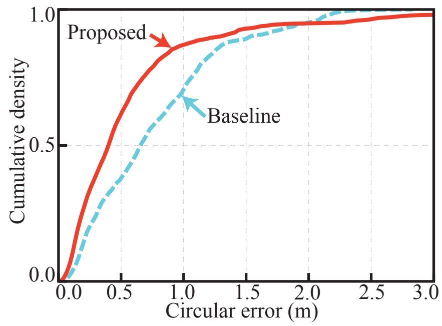

Fig. 6. Empirical cumulative distribution function (eCDF) of the localization errors for all data.

TABLE II
ERROR VALUES OF THE EXPERIMENTAL RESULTS

	Circular error (m)			Spherical error (m)		
Method	Median	Mean	90%tile	Median	Mean	90%tile
Proposed	0.40	0.58	1.28	0.42	0.60	1.29
Baseline	0.65	0.76	1.53	0.67	0.79	1.56

IV. RESULTS AND DISCUSSIONS

The evaluation results of the proposed and baseline for all data are shown in Fig. 6 and Table II. The proposed method has a relatively smaller median, mean, and CE90 (90 percentile of the localization error). However, the proposed method has large errors where the error exceeds 2 m. These samples are caused during the initialization process of the localization, or when receiving UWB signal after being unable to observe for a while. In these conditions, EKF has relatively higher covariance P_A^R than when the UWB update is continuously done. It makes Kalman gain a higher value, thus susceptible to the noise of the UWB ranging and AoA. On the other hand, the baseline method is determined by off-line processing, and there is no correction during the process. As the baseline is not a real-time method, it does not contain an initialization process, thus no huge error at the beginning. The baseline method contains drift or misalignment errors, as shown below.

Fig. 7 shows an example of the localization and its error and UWB ranging. The proposed method in Fig 7(a). The baseline method has a lower error at the start and end of the path, while it has a higher error at Exhibit 3 or the left side of Exhibit 1. This is because the baseline only aligns with the coordinates at the start. The farther away from the starting area, the larger the position error because of the orientation misalignment.

As shown in Fig. 7(c)–(d), the UWB measurements are continuously available when moving between the exhibits. This is because the user holds the navigation robot in the left hand in this term, and the UWB tag is very close to the phone. In these terms, the phone-side measurement has several ramps, while tag-side measurements are stable under 1 m. This might be caused by the phone's internal filtering. If the user is in the exhibit and leaves the navigation robot, UWB measurements are occasionally available in Exhibit 2 and Exhibit 3, but they are not available in Exhibit 1. This is because, in Exhibit 2 and Exhibit 3, the user occasionally passes through an area with good LOS and nearby from the exits of the exhibits. Exhibit 1 has a relatively large space, and its entrance and exit of the exhibit are approximately 20 m apart. This prevents measuring the UWB signal while walking inside Exhibit 1 while the navigation robot waits at the exit.

There is a place in Exhibit 3 where a significant error occurs near time 400 sec in Fig. 7(b). This is when the distance between the user and the UWB is shortened by the person coming out from the inside of the metal exhibit until around 400 seconds, and then a significant error occurs when the UWB moves to the upper left of Fig. 7(a). Since the navigation robot is waiting at the entrance of Exhibit 3, the UWB may be blocked by the user's body, making measurement hard while inside the exhibit. This term may accumulate covariance of the estimate P_A^R, making it sensitive to UWB signal update. Similar errors are observed during Exhibit 2 (approx. 200 sec) and when exiting Exhibit 3 (approx. 480 sec) and Exhibit 1 (approx. 740 sec). In future work, it is necessary to make the Kalman gain not too large for the first UWB update in a long time.

Fig. 8 shows UWB reception status for all data. In locations where it moves with the navigation robot, measurements are basically obtained in LOS. In locations away from the navigation robot, UWB can be measured for a while after separation, and there are both LOS/NLOS decisions depending on the situation. On the other hand, not many UWB measurements are obtained in locations where the robot is hidden in the shadow of a structure, as in the upper left of Exhibit 3, or in locations where it is generally 10 m away, as in the bottom side of Exhibit 1. Therefore, when viewing any of the exhibits, it is necessary to perform dead reckoning, etc. for some time when walking away from the navigation robot, even if there is no correction by the UWB. Since the proposed method can estimate the position and posture to some extent in the section where the UWB can measure, the position error is generally constant even if the robot goes to a place where the UWB cannot measure, as shown in Fig 7(a)–(b). When developing other positioning methods as future works, it is necessary to propose methods based on the assumption that the UWB will not be available a few meters away from the robot.

V. CONCLUSION

This paper proposes a positioning method advantageous for navigation for people with visual impairments. Utilizing the navigation robot's pose and relative position information aims to improve the navigation experience for people with visual impairments, who would otherwise have difficulty achieving navigation using only a smartphone. The proposed method was evaluated using data collected in a museum with actual exhibits and paths for six participants, including two participants with visual impairments. As a result, the proposed method achieves a median error of 0.40 m across all participants, while the baseline method, which uses ground truth at the start, was 0.65 m.

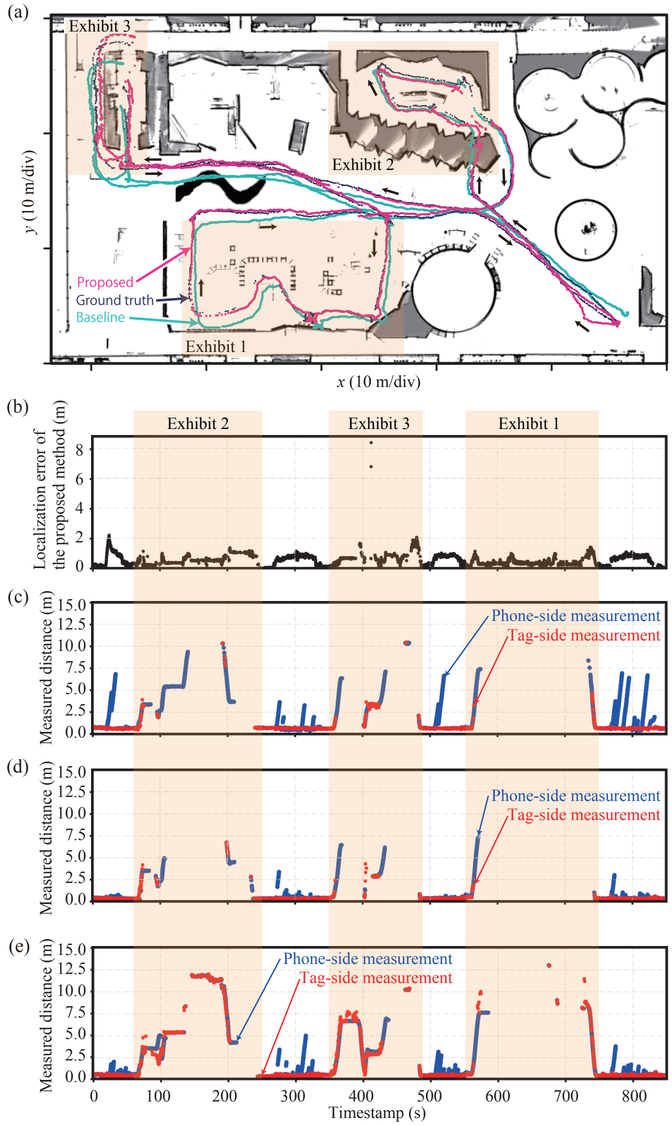

Fig. 7. Example of the localization for a path. (a) Example of localization results by the proposed method and the baseline method, (b) corresponding localization error of the proposed method, (c) corresponding timeline of the measured distance by UWB tag 1, (d) UWB tag 2, and (e) UWB tag 3.

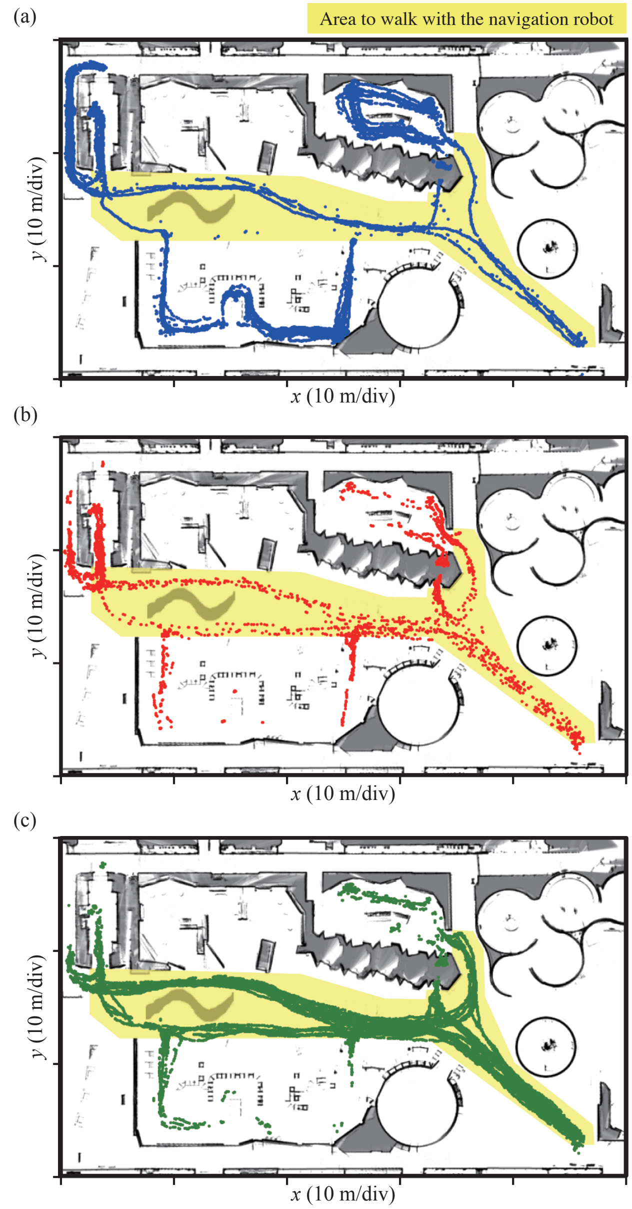

Fig. 8. UWB signal reception status for all data. (a) Locations where no UWB signals were received, (b) locations where UWB signals are received and classified as NLOS, (c) locations where UWB signals are received and classified as LOS

ACKNOWLEDGMENT

This paper is based on results obtained from a project, Programs for Bridging the gap between R&D and the IDeal society (society 5.0) and Generating Economic and social value (BRIDGE)/Practical Global Research in the AI × Robotics Services, implemented by the Cabinet Office, Government of Japan. The authors thank the participants, the staffs at Miraikan (The National Museum of Emerging Science and Innovation) and the staffs at AIST for their cooperation in this experiment.

REFERENCES

[1] D. Sato, U. Oh, K. Naito, H. Takagi, K. Kitani, and C. Asakawa, "NavCog3: An Evaluation of a Smartphone-Based Blind Indoor Navigation Assistant with Semantic Features in a Large-Scale Environment," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. Baltimore Maryland USA: ACM, Oct. 2017, pp. 270–279.

[2] C. Yoon, R. Louie, J. Ryan, M. Vu, H. Bang, W. Derksen, and P. Ruvolo, "Leveraging Augmented Reality to Create Apps for People with Visual Disabilities: A Case Study in Indoor Navigation," in *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '19. New York, NY, USA: Association for Computing Machinery, Oct. 2019, pp. 210–221.

[3] M. Kourogi and T. Kurata, "Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* Tokyo, Japan: IEEE Comput. Soc, 2003, pp. 103–112.

[4] M. Murata, D. Ahmetovic, D. Sato, H. Takagi, K. M. Kitani, and

C. Asakawa, "Smartphone-based localization for blind navigation in building-scale indoor environments," *Pervasive and Mobile Computing*, vol. 57, pp. 14–32, 2019.
[5] Y. Mashiba, H. E. B. Salih, N. Wakatsuki, K. Mizutani, and K. Zempo, "QR Code Without Impairing the Scenery and Detection System for the visually impaired," in *2020 IEEE 9th Global Conference on Consumer Electronics (GCCE)*. Kobe, Japan: IEEE, Oct. 2020, pp. 888–892.
[6] S. Ogiso, I. Mori, T. Miura, S. Nakae, T. Okuma, Y. Haga, S. Hatakeyama, K. Kimura, A. Kimura, and T. Kurata, "Integration of BLE-based proximity detection with particle filter for day-long stability in workplaces," in *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. Monterey, CA, USA: IEEE, Apr. 2023, pp. 1060–1065.
[7] D. Ahmetovic, C. Bernareggi, K. Keller, and S. Mascetti, "MusA: Artwork accessibility through augmented reality for people with low vision," in *Proceedings of the 18th International Web for All Conference*, ser. W4A '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 9–pages.
[8] A. J. Spiers and A. M. Dollar, "Design and Evaluation of Shape-Changing Haptic Interfaces for Pedestrian Navigation Assistance," *IEEE Transactions on Haptics*, vol. 10, no. 1, pp. 17–28, Jan. 2017.
[9] N. Zhou, L. Lau, R. Bai, and T. Moore, "A Robust Detection and Optimization Approach for Delayed Measurements in UWB Particle-Filter-Based Indoor Positioning," *NAVIGATION: Journal of the Institute of Navigation*, vol. 69, no. 2, p. navi.514, 2022.
[10] S. Oishi, Y. Kawamata, M. Yokozuka, K. Koide, A. Banno, and J. Miura, "C*: Cross-Modal Simultaneous Tracking and Rendering for 6-DoF Monocular Camera Localization Beyond Modalities," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5229–5236, Oct. 2020.
[11] M. Avila Soto, M. Funk, M. Hoppe, R. Boldt, K. Wolf, and N. Henze, "DroneNavigator: Using Leashed and Free-Floating Quadcopters to Navigate Visually Impaired Travelers," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. Baltimore Maryland USA: ACM, Oct. 2017, pp. 300–304.
[12] S. Kayukawa, D. Sato, M. Murata, T. Ishihara, H. Takagi, S. Morishima, and C. Asakawa, "Enhancing Blind Visitor's Autonomy in a Science Museum Using an Autonomous Navigation Robot," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg Germany: ACM, Apr. 2023, pp. 1–14.
[13] S. Hood, K. Benson, P. Hamod, D. Madison, J. M. O'Kane, and I. Rekleitis, "Bird's eye view: Cooperative exploration by UGV and UAV," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, June 2017, pp. 247–255.
[14] D. Akhihiero, U. Olawoye, S. Das, and J. Gross, "Cooperative Localization for GNSS-Denied Subterranean Navigation: A UAV–UGV Team Approach," *NAVIGATION: Journal of the Institute of Navigation*, vol. 71, no. 4, p. navi.677, 2024.
[15] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, June 2014.
[16] R. Ichikari, K. Kaji, R. Shimomura, M. Kourogi, T. Okuma, and T. Kurata, "Off-Site Indoor Localization Competitions Based on Measured Data in a Warehouse," *Sensors*, vol. 19, no. 4, pp. 763(29–pages), 2019.
[17] T. Yoshida, K. Kaji, S. Ogiso, R. Ichikari, H. Uchiyama, T. Kurata, and N. Kawaguchi, "A Survey of Ground Truth Measurement Systems for Indoor Positioning," *Journal of Information Processing*, vol. 31, no. 0, pp. 15–20, 2023.