Tag Localization by Handheld UHF RFID Reader with Optical and RFID Landmarks

Aristidis Raptopoulos Chatzistefanou, George Sergiadis, *Member, IEEE,* Antonis G. Dimitriou, *Senior, IEEE*

***Abstract*—In this paper a prototype method for manual inventorying and real-time 3D localization of RFID tagged products by a low-cost handheld UHF RFID reader is introduced. The device is equipped with a UHF RFID reader and an optical camera. "ArUco" markers with RFID tags are placed at known positions. Those optical and RF landmarks are used for the estimation of the pose of the handheld reader over time, by fusing the corresponding measurements in a Kalman filter. The reader also measures RFID tagged items at unknown locations. The positions of the tags are estimated by unwrapping the collected phase measurements, combined with the estimated antenna trajectory. Experimental results show that increased accuracy can be achieved by increasing the spatial density of the optical and RFID markers. The proposed method is expected to enhance low-cost manual inventorying, delivering 3D localization of the tagged products.**

***Index Terms*—RFID, portable, phase, localization, optical.**

I. Introduction

Radio Frequency Identification (RFID) technology has penetrated assisted living, security market, and health care in the last few years. The most important applications however reside in the field of item-level tracking and inventorying. An RFID reader outputs the existence of a tag, but not its position in space. Thanks to the improvement of the tag-IC's sensitivity, passive UHF RFID tags can be identified from distances greater than 10m from the reader's antenna. The improved read-range, though beneficial for successful inventorying, has grown the uncertainty on the tag's actual position, given that only the location of the reader can be considered known. Real-time tag/product positioning greatly enhances store/warehouse management, since it provides real-time feedback for misplaced items, expiring dates etc. As a result, in the past decade, researchers have focused on enhancing traditional inventorying with accurate localization of the inventoried tagged products. This has been addressed successfully by robots, thanks to sensory-equipment, installed on robots, which allow them to precisely calculate their poses (position and direction) over time. The robots are equipped with multiple reader-antennas which continuously interrogate all surrounding tags, collecting phase and power measurements and forming virtual antenna arrays; i.e. synthetic apertures. Processing of such data allows for accurate 3D localization of RFID-tagged items.

Manuscript received November 15, 2022

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T2EDK-02000).

All authors are with the School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece, e-mail: antodimi@auth.gr.

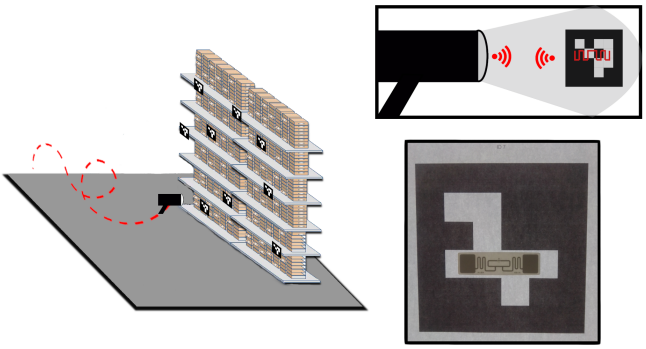


Fig. 1. The handheld RFID reader is equipped with a camera. The proposed method estimates the trace of the handheld reader by exploiting optical markers placed at known positions with phase measurements from RFID tags attached to the markers.

Due to the related cost of such robotic solutions, they are usually applicable in large retail stores and warehouses. Inventorying applications for smaller stores or setups are limited to the use of low-cost affordable equipment; i.e. commercial handheld UHF RFID readers. These readers achieve the identification (i.e. existence) of a product, but not the estimation of its position. This paper aims to enable inventorying and accurate real-time localization of tagged products by a low-cost handheld UHF RFID reader.

In this paper, we design a handheld UHF RFID reader which achieves localization of tags by combining: i) optical detection of markers at known locations, ii) phase of arrival (POA) measurements from reference RFID tags at known locations, and iii) unwrapped POA measurements from the tags to be located. A picture describing the proposed method's operational principle is shown in Fig. 1. The handheld device exploits optical-RFID tagged landmarks. Each landmark acts both as an optical marker and a reference RFID tag. The camera of the handheld device recognizes the optical marker, while the RFID reader measures the phase of the RFID tag's signal. The tag is attached on the optical marker. Landmarks are scattered among the RFID tags to be located. As the device is moved around by the user, its trajectory is estimated by combining measurement collected by the landmarks. The estimated trajectory and phase measurements are used to locate the RFID tags. The proposed method represents an extension of our work, presented in [1], where only optical markers were used for the estimation of the reader's trajectory.

We aim to improve our previous work regarding the handheld device presented in [2], where we proposed and

constructed a handheld ultra high frequency (UHF) RFID reader equipped with a 9 degrees of freedom (9dof) inertial measurement unit (IMU). The operator of the device was asked to move the device in specific manner to collect phase measurements from a **single tag** and yaw rotation measurements from the IMU. The collected data was fused using a particle filter algorithm, providing as a result an estimation of the direction and the distance to the target tag. In this paper, we combine measurements from optical and RFID landmarks to estimate the trajectory of the antenna, and then use this trajectory and unwrapped phase measurements from **multiple** tags to locate them.

Robots are often chosen to perform inventorying, since they render the process unsupervised and automated. In most applications, it is required that the robot calculates its own position. One way to achieve this is to use reference RFID tags at known locations. In [3] and [4], data from odometry sensors and phase measurements from reference RFID tags are fused using a Kalman filter to calculate the robot's pose. At the same time, the positions of tags in the environment are estimated, by processing phase measurements. In [5] and [6], data from similar sources is collected, but exploited with a different filtering approach. In [7], odomerty readings along with phase measurements and readability of reference tags are fused with a particle filter algorithm. In [8], received signal strength indicator (RSSI) measurements are exploited along with the aforementioned quantities. In [9], just a single reference tag is used. In all those cases, the 2D position of the robot is estimated, in contrast to our case in which the 3D trajectory of the antenna of the handheld device is desired.

One of the most effective ways of getting the pose of a moving robot is to consider data from light detection and ranging (lidar) and odometry sensors to perform simultaneous localization and mapping (SLAM) [10]. State of the art techniques achieve pose estimation with millimeter accuracy [11], [12]. There have been attempts to use such technology for the calculation of the 3D position of a handheld device [13]. However, expensive equipment such as a lidar and an IMU is required.

Position estimation through optical means is a promising low-cost alternative regarding the problem of calculating the trajectory of the mobile device. In [14], a camera is used to estimate the 2D trajectory of two antennas moving on a cart on a straight line. The recovered trajectory is then used to apply a direction of arrival (DoA) tag localization method. In [15], a handheld device equipped with a tracking and a depth camera is used to create a map of the environment, and then calculate its position on that map. Regarding the problem of warehouse inventorying, however, such a map-dependant solution would have reduced performance, due to the ever changing environment of the warehouse. In [16], an inverted version of our problem is examined. A static camera and antenna monitor a moving tag, trying to estimate its 2D position. In [17], a single reference tag and optical marker are used to solve the so-called kidnapped robot problem [18]. In [19], a concept of a handheld device consisting of a camera and an RFID frequency-modulated continuous-wave (FMCW) reader is presented, aiming to calculate the position of active RFID tags relatively to the handheld device's location. In [20], the distance of a handheld RFID from multiple reference tag array planes is calculated exploiting RSSI measurements. In [21], phase and IMU measurements are used to estimate the direction of the target tags from the handheld reader. In [22], a device consisting of a software-defined radio (SDR) module, two antennas, a servo motor, and a laser pointer is presented. The device operates in an environment where an RFID reader communicates with RFID tags, and calculates the angle of arrival (AoA) of the tags' signal.

As for the tag localization, most techniques rely mainly on the accurate trajectory estimation of the measuring antennas and the corresponding phase measurements. In [23] and [24] a lidar equipped robot collects phase measurements and, having unwrapped them, a synthetic-aperture radar (SAR) technique is used to create an easily solvable minimization problem to get the tags' location estimation. In [25], phase unwrapping is performed to calculate the hyperbolas on which lies the tag, the intersection of which is the tag location estimation. In [26], SAR based localization results for multiple antenna trajectories are combined to get a better position estimation of the target tags.

RSSI localization methods are not as popular among researchers, since RSSI measurements are highly ambiguous, due to the large variability of the resultant field, because of multipath. In [27], localization is divided in an offline and an online phase, and is achieved by measuring RSSI from reference tags. In [28], non-line of sight scenarios in multipath-rich environments are examined. The authors claim to exploit RSSI measurements affected by multipath in order to locate the target tags. In [29], both PoA and RSSI measurements are used to train a convolutional neural networks that performs tag localization.

Compared to previous works, the proposed method achieves accurate 3D trajectory estimation using inexpensive commercial off-the-shelf (COTS) equipment. The proposed device is operated by a human, eliminating the requirement of a robot. The device's design is as simple as possible, since it is equipped with a single antenna. Finally, accurate tag localization is achieved with low computational requirements.

In section II, the proposed method is described. Conducted experiments and their results are presented in section III. Finally, in section IV, conclusions and future work is discussed.

II. Proposed Method

The proposed method exploits measurements collected from a mobile handheld device. The measurement related components of the handheld device are:

- An RFID reader.
- An optical camera.

The data collected by the camera and the reader are used to estimate the camera's pose relatively to a global coordinate system. This is achieved by placing optical ArUco markers [30] at known positions. The optical markers are combined with reference-RFID tags, attached on top of them, in order to improve the camera's pose estimation as shown in Fig. 1. The RFID reader collects PoA measurements from both

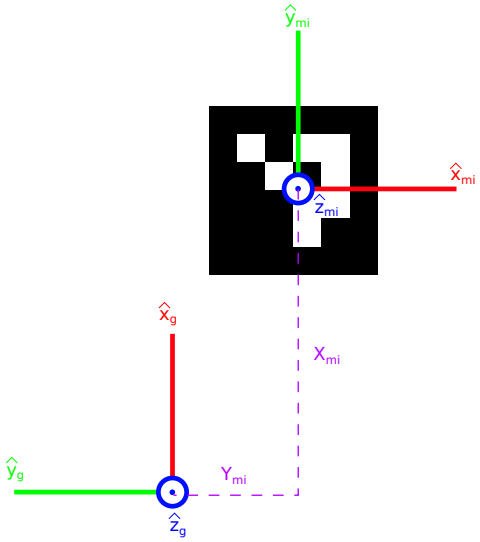


Fig. 2. An example of the local ArUco-centered and global coordinate system relation, as explained in section II-A. The basis $\left[\hat{\mathbf{x}}_m^i, \hat{\mathbf{y}}_m^i, \hat{\mathbf{z}}_m^i\right]$ is rotated and translated compared to the global basis $\left[\hat{\mathbf{x}}_g, \hat{\mathbf{y}}_g, \hat{\mathbf{y}}_g\right]$. The rotated local basis vectors can be expressed as linear combinations of the global basis vectors: $\hat{\mathbf{x}}_m^i = -\hat{\mathbf{y}}_g$, $\hat{\mathbf{y}}_m^i = \hat{\mathbf{x}}_g$, and $\hat{\mathbf{z}}_m^i = \hat{\mathbf{z}}_g$. As for the translation, the X_m^i and Y_m^i coordinates of the ArUco center are marked with purple.

the reference tags and the target-tags at unknown positions. The measurements originating from the first are fused with the camera measurements to estimate the reader's antenna trajectory. Those from the latter are processed to estimate the corresponding tag's position, given the antenna's estimated trajectory.

A. Optical Antenna Trajectory Estimation

In this section, we describe how the frames collected by the camera are used to estimate the camera's pose. Computer vision libraries from OpenCV [31] are exploited. Initially, calibration of the camera's intrinsic matrix $\mathbf{A}$, and the camera's distortion vector $\mathbf{d}_{is}$ is performed. Given specific camera settings, the calculation of $\mathbf{A}$ and $\mathbf{d}_{is}$ are performed only once.

Then, the optical marker reference system is defined. ArUco markers are used. Each marker has a unique ID recognizable by the computer vision algorithms. Let the ID of the i-th marker be ID_i. For each marker, the position of its center $\mathbf{P}_m^i = \left[X_m^i, Y_m^i, Z_m^i\right]$, its size s_i, and its orientation are defined in the global coordinate system. AruCo marker-i orientation is defined by a local orthonormal coordinate system $\left[\hat{\mathbf{x}}_m^i, \hat{\mathbf{y}}_m^i, \hat{\mathbf{z}}_m^i\right]$, so these vectors are expressed as linear combinations of the global coordinate system orthonormal basis $\left[\hat{\mathbf{x}}_g, \hat{\mathbf{y}}_g, \hat{\mathbf{z}}_g\right]$:

$$\hat{\mathbf{x}}_m^i = a_{ix}\hat{\mathbf{x}}_g + a_{iy}\hat{\mathbf{y}}_g + a_{iz}\hat{\mathbf{z}}_g \tag{1}$$

$$\hat{\mathbf{y}}_m^i = b_{ix}\hat{\mathbf{x}}_g + b_{iy}\hat{\mathbf{y}}_g + b_{iz}\hat{\mathbf{z}}_g \tag{2}$$

$$\hat{\mathbf{z}}_m^i = c_{ix}\hat{\mathbf{x}}_g + c_{iy}\hat{\mathbf{y}}_g + c_{iz}\hat{\mathbf{z}}_g \tag{3}$$

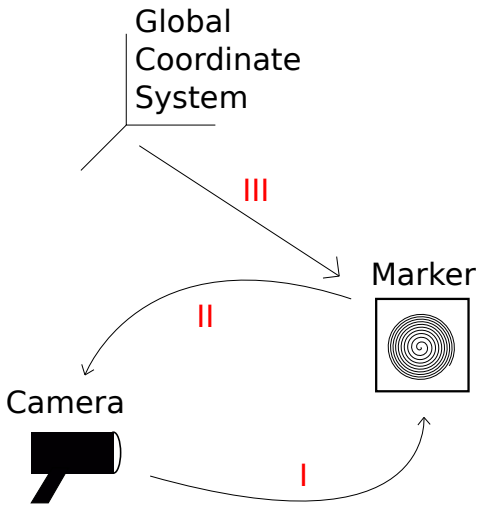


Fig. 3. An illustration of the process described in this section II-A: (I) The camera detects a marker. (II) The relative position of the camera is calculated relatively to the marker. (III) Since the pose of the marker is known, the position of the camera in the global coordinate system can be calculated.

An illustration of the above is shown in Fig. 2. Frames are continuously captured, in order to estimate the trajectory of the camera. Let $frame_n$ be the frame captured at time t_n^f. Assuming that ID_i was detected in $frame_n$, given the camera parameters $\mathbf{A}$ and $\mathbf{d}_{is}$, and the marker size s_i, the rotation vector $\mathbf{rot}_n^i$ and the translation vector $\mathbf{tr}_n^i = [trX_n^i, trY_n^i, trZ_n^i]$ can be calculated. These vectors show that if the camera is translated according to $\mathbf{tr}_n^i$, and then rotated according to $\mathbf{rot}_n^i$, it will be aligned with the i-th marker. So, if this process is reversed, the position of the camera is retrieved. First, the basis vectors of the i-th marker are rotated according to $-\mathbf{rot}_n^i$, using Rodrigues' rotation formula:

$$\mathbf{v}_{rot} = \mathbf{v}cos(\theta) + (\mathbf{k} \times \mathbf{v})sin(\theta) + \mathbf{k}(\mathbf{k} \cdot \mathbf{v})\left(1 - cos(\theta)\right) \tag{4}$$

$\mathbf{v}$ is the vector to be rotated which in our case are coordinate system unit vectors, $\mathbf{k}$ is the unit vector showing the axis of rotation, θ is the rotation angle, and $\mathbf{v}_{rot}$ is the rotated vector. In this case:

$$\mathbf{k} = \frac{-\mathbf{rot}_n^i}{|\mathbf{rot}_n^i|} \tag{5}$$

$$\theta = |\mathbf{rot}_n^i| \tag{6}$$

where $|\cdot|$ denotes the calculation of the Euclidean norm. Using (4), $\left[\hat{\mathbf{x}}_m^i, \hat{\mathbf{y}}_m^i, \hat{\mathbf{z}}_m^i\right]$ are rotated according to $-\mathbf{rot}_n^i$, and we get $\left[\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n, \left(\hat{\mathbf{y}}_m^i\right)_{rot}^n, \left(\hat{\mathbf{z}}_m^i\right)_{rot}^n\right]$. These vectors can be expressed as linear combinations of $\left[\hat{\mathbf{x}}_m^i, \hat{\mathbf{y}}_m^i, \hat{\mathbf{z}}_m^i\right]$. Using (4):

$$\begin{aligned}\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n = \; & \hat{\mathbf{x}}_m^i cos(|\mathbf{rot}_n^i|) \\ & + \left(\frac{-\mathbf{rot}_n^i}{|\mathbf{rot}_n^i|} \times \hat{\mathbf{x}}_m^i\right) sin(|\mathbf{rot}_n^i|) \\ & + \frac{-\mathbf{rot}_n^i}{|\mathbf{rot}_n^i|}\left(\frac{-\mathbf{rot}_n^i}{|\mathbf{rot}_n^i|} \cdot \hat{\mathbf{x}}_m^i\right)\left(1 - cos(|\mathbf{rot}_n^i|)\right)\end{aligned} \tag{7}$$

There are no unknown variables in the right part of (7). So, $\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n$ can be calculated and expressed as linear combination of $\left[\hat{\mathbf{x}}_m^i, \hat{\mathbf{y}}_m^i, \hat{\mathbf{z}}_m^i\right]$:

$$\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n = (a_{ix})'_n\hat{\mathbf{x}}_m^i + (a_{iy})'_n\hat{\mathbf{y}}_m^i + (a_{iz})'_n\hat{\mathbf{z}}_m^i \tag{8}$$

Similarly for $\left(\hat{\mathbf{y}}_m^i\right)_{rot}^n$ and $\left(\hat{\mathbf{z}}_m^i\right)_{rot}^n$:

$$\left(\hat{\mathbf{y}}_m^i\right)_{rot}^n = (b_{ix})'_n\hat{\mathbf{x}}_m^i + (b_{iy})'_n\hat{\mathbf{y}}_m^i + (b_{iz})'_n\hat{\mathbf{z}}_m^i \tag{9}$$

$$\left(\hat{\mathbf{z}}_m^i\right)_{rot}^n = (c_{ix})'_n\hat{\mathbf{x}}_m^i + (c_{iy})'_n\hat{\mathbf{y}}_m^i + (c_{iz})'_n\hat{\mathbf{z}}_m^i \tag{10}$$

The position of the camera relatively to the i-th marker $\mathbf{Pc}_n^i$ is calculated, if a translation is performed according to $-\mathbf{tr}_n^i$, using the basis $\left[\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n, \left(\hat{\mathbf{y}}_m^i\right)_{rot}^n, \left(\hat{\mathbf{z}}_m^i\right)_{rot}^n\right]$:

$$\mathbf{Pc}_n^i = -trX_n^i\left(\hat{\mathbf{x}}_m^i\right)_{rot}^n - trY_n^i\left(\hat{\mathbf{y}}_m^i\right)_{rot}^n - trZ_n^i\left(\hat{\mathbf{z}}_m^i\right)_{rot}^n \tag{11}$$

From (1)-(3) and (8)-(11) we get:

$$\mathbf{Pc}_n^i = \mathbf{Tr}_n^i(\mathbf{M}_i)'_n\mathbf{M}_i\begin{bmatrix}\hat{\mathbf{x}}_g\\ \hat{\mathbf{y}}_g\\ \hat{\mathbf{z}}_g\end{bmatrix} \tag{12}$$

$$\mathbf{Tr}_n^i = \begin{bmatrix}-trX_n^i & 0 & 0\\ 0 & -trY_n^i & 0\\ 0 & 0 & -trZ_n^i\end{bmatrix} \tag{13}$$

$$(\mathbf{M}_i)'_n = \begin{bmatrix}(a_{ix})'_n & (a_{iy})'_n & (a_{iz})'_n\\ (b_{ix})'_n & (b_{iy})'_n & (b_{iz})'_n\\ (c_{ix})'_n & (c_{iy})'_n & (c_{iz})'_n\end{bmatrix} \tag{14}$$

$$\mathbf{M}_i = \begin{bmatrix}a_{ix} & a_{iy} & a_{iz}\\ b_{ix} & b_{iy} & b_{iz}\\ c_{ix} & c_{iy} & c_{iz}\end{bmatrix} \tag{15}$$

It is reminded that $\mathbf{Pc}_n^i$ is the position of the camera relatively to the i-th marker. The marker's position $\mathbf{P}_m^i$ is added to $\mathbf{Pc}_n^i$, in order to calculate the corresponding position on the global coordinate system $\mathbf{Pg}_n^i$, :

$$\mathbf{Pg}_n^i = \left(\mathbf{Tr}_n^i(\mathbf{M}_i)'_n\mathbf{M}_i + \begin{bmatrix}X_m^i & 0 & 0\\ 0 & Y_m^i & 0\\ 0 & 0 & Z_m^i\end{bmatrix}\right)\begin{bmatrix}\hat{\mathbf{x}}_g\\ \hat{\mathbf{y}}_g\\ \hat{\mathbf{z}}_g\end{bmatrix} \tag{16}$$

An illustration of the process described in this section is shown in Fig. 3: The camera detects a marker. The position of the camera is calculated relatively to the marker. Since the pose of the marker is known, the position of the camera in the global coordinate system can be calculated. At the end of this process, the camera's position $\mathbf{Pg}_n^i$ is calculated based on the detected i-th marker at time t_n^f. A captured frame and the detected markers are shown in Fig. 4.

Fig. 4. An example of the detection of optical markers. The corners of the black squares are detected, and by the combination of the black and white blocks inside the square the orientation of the marker is extracted, as well as its id. The local coordinate system of each detected marker is also drawn on the captured frame.

B. Optical Position Estimation Filtering

The aforementioned method results in numerous camera-position estimations, since each identified ArUco marker per frame corresponds to an estimated camera position. For instance, when three markers are identified in a single frame, the output would be three "possible" positions, which do not necessarily coincide, due to poor quality images; e.g. poor focus, long distance, greatly tilted camera, etc.. Thanks to the frame-rate, which depends on the selected camera (in our case 30Hz), the algorithm would end up with more than 100 position-estimations per s.

In this section, the process of discarding poor-quality images and exploiting the remaining ones is described. Low quality captured frames lead to position estimations with great errors. Every estimation is required to meet some conditions in order to be considered valid. The first condition is that position estimation $\mathbf{Pg}_n^i$ must be in front of the corresponding i-th marker. The marker is only visible from one side, so the frame on which the i-th marker was detected was captured while the camera was facing the visible side of the marker. This condition can be expressed as:

$$\overline{\mathbf{P}_m^i\mathbf{Pg}_n^i}\cdot\hat{\mathbf{z}}_m^i > 0 \tag{17}$$

where $\overline{\mathbf{P}_m^i\mathbf{Pg}_n^i}$ is the vector connecting the center of the i-th marker with the estimated camera position. $\hat{\mathbf{z}}_m^i$ is the normal vector on the visible side of the marker. So, if the projection of $\overline{\mathbf{P}_m^i\mathbf{Pg}_n^i}$ on $\hat{\mathbf{z}}_m^i$ is positive, $\mathbf{Pg}_n^i$ is in the "correct" half-space.

The second condition is related to the distance of the estimated position $\mathbf{Pg}_n^i$ to the center $\mathbf{P}_m^i$ of the i-th marker. As the distance between the camera and the marker increases, the quality of the marker's depiction on the captured frames decreases. Let R_{max} be the maximum distance between the camera and the marker for which we consider that the captured frame can provide an accurate estimation of the position of the camera. On the other hand, the marker is required to be entirely visible on the captured frame in order to be detected. So, the camera cannot be arbitrarily close to the marker. Let

the minimum distance be R_{min}. The second condition can be expressed as:

$$R_{min} < |(\mathbf{P}^i_m \mathbf{P}\mathbf{g}^i_n)| < R_{max} \tag{18}$$

The third and final condition ensures that the marker is detected only when seen from a limited angular range. $|(\mathbf{P}^i_m \mathbf{P}\mathbf{g}^i_n)|$ is the Euclidean norm of vector $\overline{\mathbf{P}^i_m \mathbf{P}\mathbf{g}^i_n}$. As the angle between the camera's facing direction and the marker's $\hat{\mathbf{z}}^i_m$ vector increases, the area occupied by the marker's depiction on the captured frame decreases. So, the ambiguity of the measurement is expected to increase. So the frames in which the i-th marker was detected from an angle greater than ang_{max} are discarded. The following condition must be met:

$$arccos\left(\frac{\overline{\mathbf{P}^i_m \mathbf{P}\mathbf{g}^i_n} \cdot \hat{\mathbf{z}}^i_m}{|(\mathbf{P}^i_m \mathbf{P}\mathbf{g}^i_n)|}\right) < ang_{max} \tag{19}$$

Let the remaining filtered measurements, after discarding those that do not fulfill (17)-(19), be $\mathbf{Pf}^i_n$. The filtering process continues by identifying overlapping time windows and grouping the measurements in each window. The length of each time window $TW_j, j = 1, 2, \ldots, J$ varies, so that each window contains at least N_{frame} measurements $\mathbf{Pf}^i_n$.

The last step of filtering aims to extract one position estimation from each time window TW_j. The position estimations $\mathbf{Pf}^i_n$ of each window are divided in clusters, using a k-means clustering algorithm. The center of the cluster with the greatest size is $\mathbf{Pw}_j$, and is matched with the midpoint tw_j of the corresponding time window TW_j. So, the estimation set that is used in the following steps of the algorithm are the J position estimations $\mathbf{Pw}_j$ and the J corresponding times tw_j. An example of how clustering is performed in each time window is shown in Fig. 5.

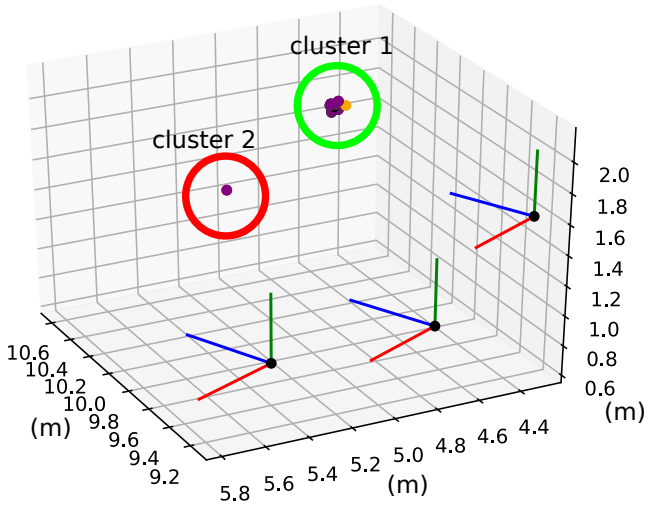


Fig. 5. The filtered estimations $\mathbf{Pf}^i_n$ of the window are divided in clusters, indicated here by the green and red circle. The center of the most populated cluster (green) represents the estimated position $\mathbf{Pw}_j$ of the specific time window. $\mathbf{Pw}_j$ is shown with a black dot. The corresponding ground truth position is shown with an orange dot. The optical landmarks and their orientation are marked with black dots and colored lines, as in Fig. 2. Axis dimensions are in meters (m).

C. Handheld Device Velocity Estimation

As discussed earlier, RFID PoA measurements are also collected from the RFID tags attached to each ArUco marker. Those measurements are fused in a Kalman filter to estimate the velocity of the handheld device as explained below. The reader collects phase measurements of the tag's backscattered signal, wrapped in $[0, \pi)$. Initially, the phase measurements are unwrapped. The measurements include noise, which is modeled as a normally distributed variable. The measured phase is:

$$\phi_{meas} = (\phi + \phi_{noise}) \ \ mod \ \ (\pi) \tag{20}$$

$$\phi = \phi_p + \phi_o \tag{21}$$

$$\phi_p = \frac{4\pi}{\lambda} d \tag{22}$$

$$\phi_{noise} \sim N(0, s_{phase}), \tag{23}$$

where the phase of the received signal ϕ is considered as the sum of two components ϕ_p and ϕ_o. ϕ_p is the phase accumulated due to electromagnetic wave propagation, ϕ_o is a phase offset due to cabling and the related hardware. ϕ_{noise} is the phase measurement noise, λ the carrier's wavelength, and d the distance from the antenna to the tag.

Let $\mathbf{P}_{ant}(t)$ be the position of the antenna at time t. Reference RFID tags are attached to the optical markers; hence share the same coordinates $\mathbf{P}^i_m$. Regarding the position of the antenna, (22) becomes:

$$\phi_p(t, i) = \frac{4\pi}{\lambda} |(\mathbf{P}^i_m \mathbf{P}_{ant}(t))| \tag{24}$$

The rate of change of (24) is:

$$\begin{aligned} \frac{\Delta}{\Delta t}(\phi_p(t, i)) &= \frac{4\pi}{\lambda} \frac{\Delta}{\Delta t} \left(|(\mathbf{P}^i_m \mathbf{P}_{ant}(t))|\right) \\ \frac{\lambda}{4\pi} \frac{\Delta}{\Delta t}(\phi_p(t, i)) &= \frac{\Delta}{\Delta t} \left(|(\mathbf{P}^i_m \mathbf{P}_{ant}(t))|\right) \end{aligned} \tag{25}$$

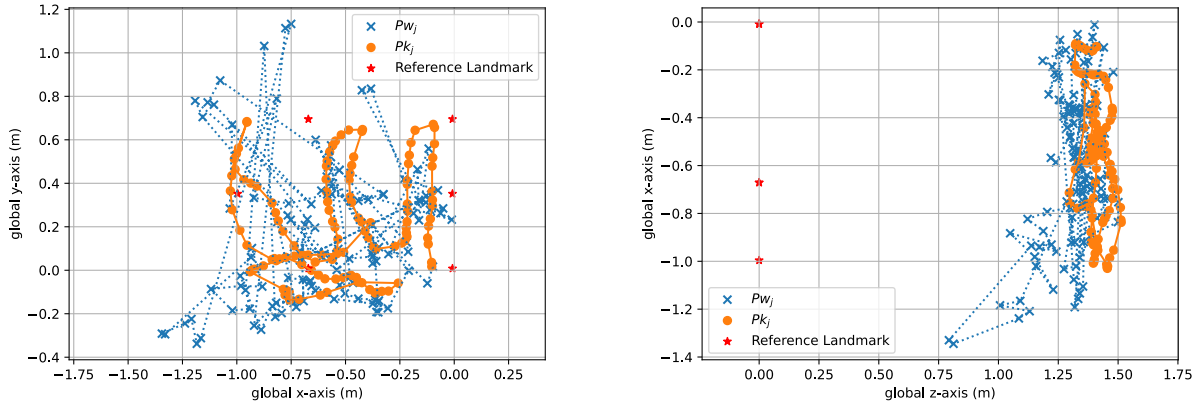
The right part of (25) can be interpreted as the magnitude of the velocity of the antenna in the direction of the i-th reference tag [32]. Let that radial velocity be $\bar{\mathbf{U}}_{rad}(t, i)$:

$$\bar{\mathbf{U}}_{rad}(t, i) = -\frac{\lambda}{4\pi} \frac{\Delta}{\Delta t}(\phi_p(t, i)) \, \hat{\mathbf{u}}_{rad}(t, i) \tag{26}$$

$$\hat{\mathbf{u}}_{rad}(t, i) = \frac{\overline{\mathbf{P}_{ant}(t) \mathbf{P}^i_m}}{|(\mathbf{P}_{ant}(t) \mathbf{P}^i_m)|} \tag{27}$$

where, $\hat{\mathbf{u}}_{rad}(t, i)$ is the radial unit vector from the position of the antenna to the i-th reference tag. Let the real velocity of the antenna be $\bar{\mathbf{U}}_{real}(t) = [U^x_{real}(t), U^y_{real}(t), U^z_{real}(t)]^T$. $\bar{\mathbf{U}}_{rad}(t, i)$ is the projection of $\bar{\mathbf{U}}_{real}(t)$ on the corresponding radial vector $\hat{\mathbf{u}}_{rad}(t, i)$:

$$\hat{\mathbf{u}}_{rad}(t, i) \cdot \bar{\mathbf{U}}_{real}(t) = -\frac{\lambda}{4\pi} \frac{\Delta}{\Delta t}(\phi_p(t, i)) \tag{28}$$



(a) x-axis and y-axis coordinates.

(b) z-axis and x-axis coordinates.

Fig. 6. An example of the measurement filtering process described in Section II-D. The x-axis and y-axis coordinates are shown in the left figure, while the z-axis and x-axis in the right one. The orange dots are the Kalman filtered estimations $\mathbf{Pk}_j$, while the blue "x"s the windowed estimations $\mathbf{Pw}_j$. The red stars represent the positions of the optical markers and reference RFID tags.

$U^x_{real}(t)$, $U^y_{real}(t)$, $U^z_{real}(t)$ in (28) are evaluated by solving the following minimization problem:

$$\bar{\mathbf{U}}_{real}(t) = \arg\min_{\bar{\mathbf{U}}} \sum_i \left(f_{cost}\left(\bar{\mathbf{U}}, i, t\right)\right) \tag{29}$$

$$f_{cost}\left(\bar{\mathbf{U}}, i, t\right) = \left(\hat{\mathbf{u}}_{rad}(t,i) \cdot \bar{\mathbf{U}} + \frac{\lambda}{4\pi}\frac{\Delta}{\Delta t}\left(\phi_p(t,i)\right)\right)^2 \tag{30}$$

An example of the 2D projection of the real velocity of the antenna on the directions of the reference tags is shown in Fig. 8.

The phase measurements are processed to derive a better estimation of the radial velocity. Initially, the measured data are unwrapped. Then, a Savitzky–Golay filter is applied to the unwrapped phase measurements, which reduces the hardware's introduced phase-noise. The length of the filter's window is in the order of seconds. A representative example of the process is shown in Fig. 7, resulted from experimental measurements. The actual rate is $-0.7186\ rad/s$ and the estimated is $-0.6606\ rad/s$; the expected error is small.

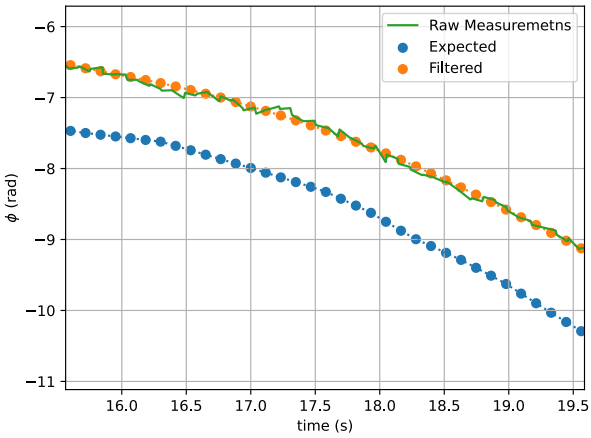


Fig. 7. Experimentally collected phase measurements, filtered as described in section II-C. A SLAM enabled robot collects phase measurements from an RFID tag at a known location. The orange line shows the filtered measurements, whereas the blue the expected phase measurements, computed according to the robot's trajectory. The green line show the raw phase measurements as collected by the reader. The filtered measurements are used to calculate the rate of change $\Delta\phi/\Delta t$ for the antenna's velocity estimation.

D. Kalman Filtering

The optical measurements from sections II-A and II-B are fused with the velocity estimation from II-C using a Kalman filter. The estimated state is the 3D position $\mathbf{s}_j$ of the antenna:

$$\mathbf{s}_j = [sx_j, sy_j, sz_j]^T \tag{31}$$

sx_j, sy_j, sz_j are the 3D coordinates of the antenna at time tw_j. The system's model is:

$$\mathbf{s}_j = F\mathbf{s}_{j-1} + B\mathbf{u}_{j-1} + \mathbf{w}_{j-1} \tag{32}$$

The filter's task is to model random movements of the handheld device. F is the state transition matrix, indicating the starting position of the previous state:

$$F = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{33}$$

B is the control matrix and $\mathbf{u}_{j-1}$ is the control vector, representing the velocity of the antenna, which represents the cause for the change of the current state. $\mathbf{u}_{j-1}$ is estimated through the process described in II-C. It is reminded that phase measurements from tags at known positions are collected and used to estimate the velocity of the antenna, assuming that its position is also known.

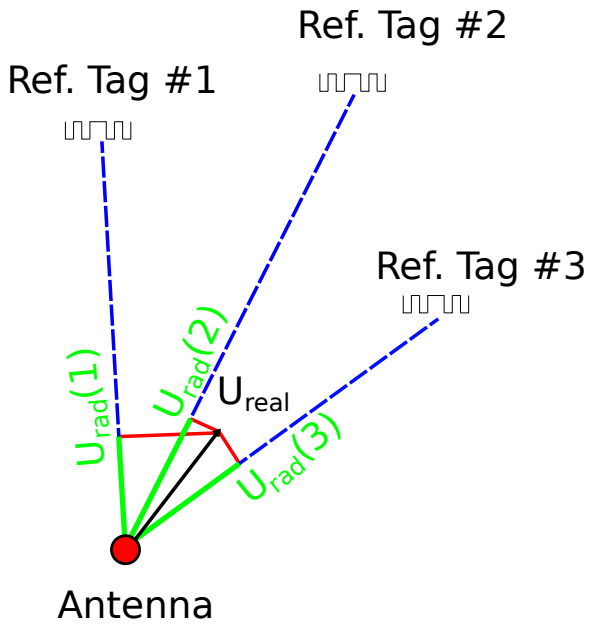


Fig. 8. An illustration of the projection of the real velocity of the antenna on the directions of the reference tags. The projections of $\bar{\mathbf{U}}_{real}$ on the known directions of the reference tags can be calculated. Knowing $\bar{\mathbf{U}}_{rad}(1)$, $\bar{\mathbf{U}}_{rad}(2)$, $\bar{\mathbf{U}}_{rad}(3)$ allows one to estimate $\bar{\mathbf{U}}_{real}$.

$$\begin{aligned} \mathbf{u}_{j-1} &= \bar{\mathbf{U}}_{real}(tw_{j-1}) \\ \mathbf{u}_{j-1} &= [U^x_{real}(tw_{j-1}), U^y_{real}(tw_{j-1}), U^z_{real}(tw_{j-1})]^T \end{aligned} \tag{34}$$

B is the time interval Δt between the state estimation $j-1$ and j:

$$B = \Delta t \tag{35}$$

Finally, $\mathbf{w}_{j-1}$ is a random variable used to model the system's ambiguities and the device's random acceleration:

$$\mathbf{w}_{j-1} \sim N(0, Q) \tag{36}$$

$$Q = s_a^2 I_3 \tag{37}$$

where s_a is the filter parameter regarding acceleration of the device. The corresponding measurement model is:

$$\mathbf{z}_j = H\mathbf{s}_j + \mathbf{v}_j \tag{38}$$

where $\mathbf{z}_j$ is the measurement, H is the measurement matrix, and $\mathbf{v}_j$ is the measurement noise vector. The measurement in our case is the optical position estimation $\mathbf{Pw}_j$ described in sections II-A and II-B. So:

$$\mathbf{z}_j = \left[\mathbf{Pw}_j \cdot \hat{\mathbf{x}}_g, \mathbf{Pw}_j \cdot \hat{\mathbf{y}}_g, \mathbf{Pw}_j \cdot \hat{\mathbf{z}}_g\right]^T \tag{39}$$

$$H = I_3 \tag{40}$$

$$\mathbf{v}_j \sim N(0, R) \tag{41}$$

$$R = s_m^2 I_3 \tag{42}$$

where s_m is the filter parameter regarding the ambiguity of the measurements. The Kalman filtering process is defined by the following equations:

$$\hat{\mathbf{s}}_j^- = F\hat{\mathbf{s}}_{j-1}^+ + B\mathbf{u}_{j-1} \tag{43}$$

$$P_j^- = FP_{j-1}^+F^T + Q \tag{44}$$

$$K_j = P_j^- H^T \left(R + HP_j^- H^T\right)^{-1} \tag{45}$$

$$\hat{\mathbf{s}}_j^+ = \hat{\mathbf{s}}_j^- + K_j\left(\mathbf{z}_j - H\hat{\mathbf{s}}_j^-\right) \tag{46}$$

$$P_j^+ = \left(I_3 - K_j H\right) P_j^- \tag{47}$$

In the above, $\hat{\mathbf{s}}_j^+$ is the estimation of the device's position at time tw_j. $\hat{\mathbf{s}}_{j-1}^+$ is used to calculate $\mathbf{u}_{j-1}$ according to section II-C. Through the resulting estimations, the final estimation of the antenna's trajectory $\mathbf{Pk}_j$ is calculated. An example of the windowed estimations $\mathbf{Pw}_j$ and the Kalman filtered estimations $\mathbf{Pk}_j$ is shown in Fig. 6.

E. Tag Localization

Given that the virtual antenna trajectory of the handheld device is estimated, as presented earlier, the final part of the proposed method estimates the locations of the unknown RFID tags. As stated in section II-C, the high read rate facilitates phase unwrapping of the collected measurements per tag. Let the unwrapped phase of tag-l at time tw_j be $\phi_u(l, tw_j)$. The 3D position of tag-l $\mathbf{P}_l$ is estimated through the following minimization problem, as in [23]:

$$\{\mathbf{P}_l, c_l\} = \arg\min_{\mathbf{P},c} \sum_j g^2_{cost}(\mathbf{P}, c, j, l) \tag{48}$$

$$g_{cost}(\mathbf{P}, c, j, l) = \frac{4\pi}{\lambda}|(\mathbf{P}\ \mathbf{Pk}_j)| - \phi_u\left(l, tw_j\right) + c \tag{49}$$

Equation (49) evaluates the similarity of the collected measurement ϕ_u at time t_j with the expected one, assuming that the position of the tag is $\mathbf{P}$. The higher the value of (49), the lower the similarity. The position that minimizes the sum of g_{cost} for all collected measurements is the estimated position of the corresponding tag, as shown in (48).

III. Experimental Results

In this section, experimental results of the proposed method are presented. The rest of the section is divided in three parts: The accuracy of the antenna trajectory estimation is examined in section III-A, and that of the tag localization in section III-B. In section III-C, a SLAM enabled robot is used to compare the localization performance in two cases: using the estimated and the ground truth trajectory of the antenna.

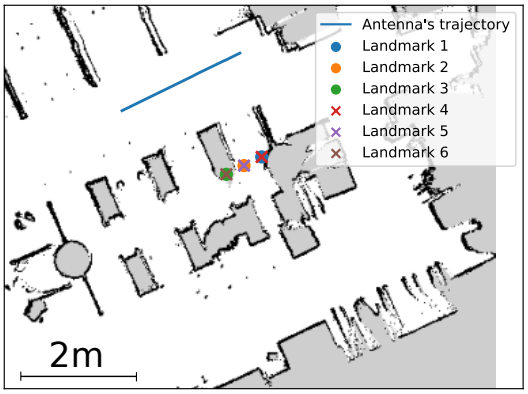


Fig. 9. The map created by the robot in one of the experiments described in section III-A, where both the trajectory of the antenna and the landmarks are shown. Three of the landmarks (with dots) are below the other three.

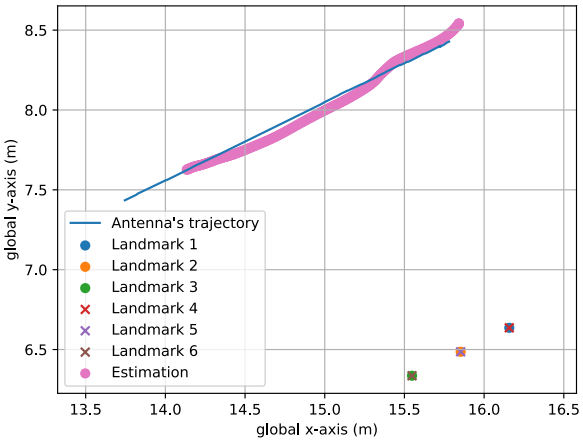


Fig. 11. Comparative results of experiment $E.A2$ as described in section III-A.

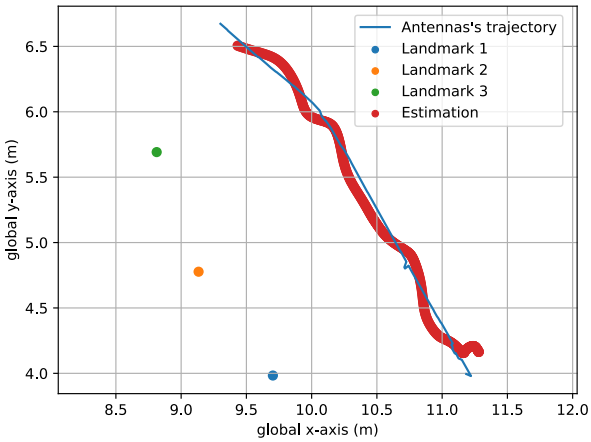


Fig. 10. The results of experiment $E.A1$ as described in section III-A. The blue line is the ground truth trajectory of the antenna, the red dots are the estimated positions $\mathbf{Pk}_j$.

A. Trajectory Estimation Accuracy

The ground truth position of the moving antenna is required, in order to evaluate the accuracy of the estimated antenna trajectory. The ground truth of the device's position is acquired by attaching the device on top of a SLAM-capable robot. The robot, carrying the antenna and the camera, is equipped with a low-noise 360^o lidar, which guarantees accurate map creation and localization of its own pose (at cm accuracy). The acquired position of the antenna on top of the robot at time t is:

$$\mathbf{P}_{ant}(t) = [x_{ant}(t), y_{ant}(t), z_{ant}(t)] \tag{50}$$

At the same time, the camera of the device (on top of the robot) detects the optical/RFID markers, thus estimating the device's positions $\mathbf{Pk}_j$ as explained in sections II-A to II-D. To evaluate the accuracy of the estimation, the error Er_T between the robot's ground-truth positions $\mathbf{P}_{ant}(tw_j)$ and the estimated positions $\mathbf{Pk}_j$ is:

$$Er_T = E\left[|(\mathbf{P}_{ant}(tw_j)\ \mathbf{Pk}_j)|\right] \tag{51}$$

where $E[\cdot]$ represents the mean value. The map created by the robot, the antenna's ground truth positions, and the positions of the landmarks in one of the experiments is shown in Fig. 9.

Two experiments are presented, each with different number of landmarks and size of optical markers. In the first experiment $E.A1$ three landmarks were used. The size of the optical markers was $18cm$ x $18cm$. In the second experiment $E.A2$ six landmarks were used, but the size of the optical markers was $9cm$ x $9cm$. Given that the same camera and configuration is used, larger optical markers are expected to provide better position estimations $\mathbf{Pw}_j$, since the projection of the optical marker on the captured frame is larger, thus providing higher Signal-to-Noise ratio (SNR).

The ground truth antenna trajectory and estimation of $E.A1$ on the xy-plane is shown in Fig. 10, along with an illustration of how the estimated positions were matched with the ground truth in order to calculate (51). The results of $E.A2$ on the xy-plane are shown in Fig. 11. The calculated 3D errors Er_T for $E.A1$ and $E.A2$ are $0.19m$ and $0.22m$ respectively, verifying a small increase of the mean error for the smaller landmarks. In general, the estimated trajectories from the proposed optical/RFID system are very close to the actual trajectories. Next, we examine the final error, on the estimation of the positions of the unknown tags from a manual operated handheld device.

B. Tag Localization Accuracy

Tagged items were placed on shelves as shown in Fig. 12. Thirty six target-tags and six landmarks were distributed in the $1.4m$ x $1.2m$ experimental area. The size of the optical markers was $9cm$ x $9cm$.

The handheld device was moved around randomly by the user who was facing the shelves standing in front of them

Fig. 12. RFID tagged optical landmarks were placed on the faces of some shelves. The handheld device was moved in front of the shelves, facing the landmarks. The trajectory of its antenna was calculated as explained in sections II-A to II-D, and was used, in order to calculate their positions.

as would be done in manual inventorying. The device was moved so that its camera could detect the optical markers and its directional antenna the RFID tags. The estimated trajectory of the antenna in one of the experiments and the locations of the target-tags are shown in Fig. 13. In order to estimate the positions of the tags the minimization problem of (48), (49) is solved.

Three experiments were conducted. Characteristic results are shown in Fig. 14. Let $\mathbf{G}_l = [G_x(l), G_y(l), G_z(l)]$ be the ground truth location of the l-th of L tags and $\mathbf{P}_l = [P_x(l), P_y(l), P_z(l)]$ its estimated location. In order to evaluate the results, for each experiment we calculate the mean error Er_{3D} between the ground truth and estimated location:

$$Er_{3D} = \frac{1}{L}\sum_{l=1}^{L}|(\mathbf{P}_l\ \mathbf{G}_l)| \tag{52}$$

The resulted Er_{3D} are summarized in Table I. In the same table, the number of antenna trajectory points and RFID tag reads for each experiment are given. These numbers are an indicator of the quality of the measured data. The number of points of the estimated antenna trajectory depends on the number of frames, where the optical markers were successfully detected. Similarly, the number of RFID-reads affects the antenna trajectory estimation and ultimately the tag-localization accuracy. From all experimental results of Table I, the mean value of Er_{3D} is $0.953m$. The corresponding 3D error is considered small for manual inventorying, bearing in mind that it is smaller than 1/10 of the reading range of the RFID reader.

Finally, we examine whether or not the required computational time grants the proposed method suitable for real-time applications. In Table II the required computational time to get the tag localization results for each experiment is shown. This time includes both the trajectory estimation of the antenna and the tag localization. Along the computational time, the duration of each experiment is presented, that is for how

TABLE I
EXPERIMENT RESULTS FOR TAG LOCALIZATION ACCURACY EVALUATION.

Experiment	Er_{3D} (m)	Trajectory Points	RFID Reads
# 1	0.93	120	12403
# 2	1.03	98	7461
# 3	0.90	131	14047

TABLE II
COMPUTATIONAL TIME FOR TAG LOCALIZATION.

Experiment	Computational Time (s)	Measurement Time (s)
# 1	20	81
# 2	12	49
# 3	20	96

long measurements were collected by the camera and the RFID reader. In all cases, the required computational time is no greater than $20s$. Additionally, the measurement time is always at least four times greater than the corresponding computational time. The results indicate high applicability of the proposed method for real-time applications.

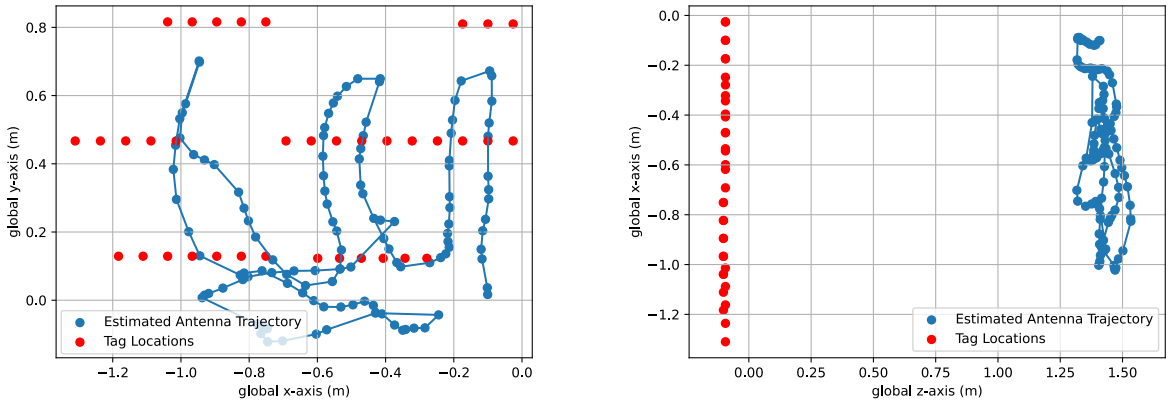
C. Effect of trajectory-estimation error on the tag localization accuracy

Additional experiments using a SLAM enabled robot were conducted to compare the localization performance in two cases: using the estimated and the ground truth trajectory of the antenna. The robot moved in an environment with high density of optical markers and reference tags, shown in Fig. 15. Apart from the reference, target-tags were placed at known locations. Having estimated the robot's trajectory using the proposed method, we performed tag-localization, using i) the estimated antenna trajectory, and ii) the ground truth antenna trajectory. For the estimation of the antenna-trajectory, we considered two cases; one with all 10 optical and RFID markers (high density estimations) and one with only three optical and RFID markers (low density estimations).

The results are presented in Fig. 16. In the low-density case, the estimated trace appears "bended" at the top of the figure with respect to the ground truth and the "high density" case, which resulted in larger localization error of the tags at the top of the figure (opposite to the bended trace). In the high-density case, the trajectory estimation error was $14cm$ with a standard deviation of $2cm$, which resulted in a mean tags' localization error of $22cm$; comparable to the ground truth mean localization error of $19cm$. When only 3 markers were used, the mean trajectory error was $16cm$; however the mean localization error was increased to $37cm$, due to the erroneously estimated curvature of the trajectory, affecting the tags at the top of the figure.

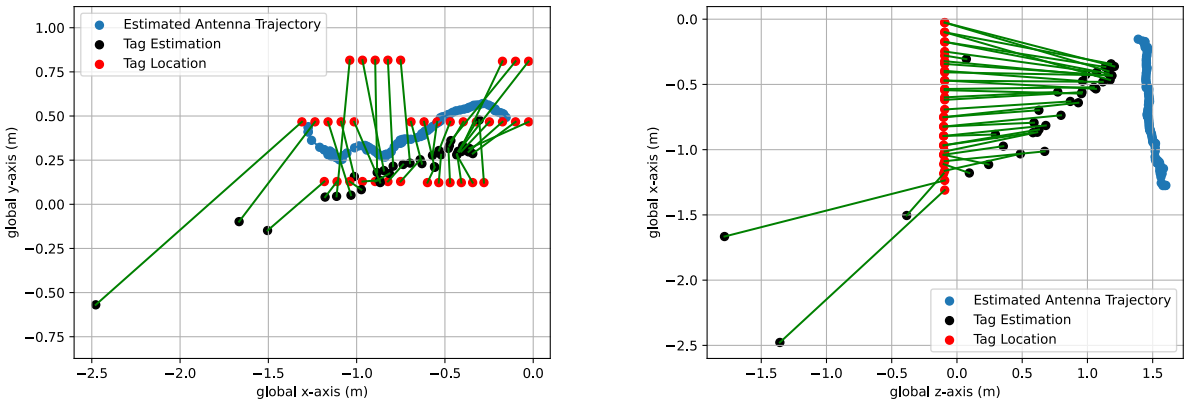
IV. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a method for RFID tag localization using a handheld device consisting of an RFID



(a) x-axis and y-axis coordinates. (b) z-axis and x-axis coordinates.

Fig. 13. An illustration of a tag localization experiment, presented in section III-B in side (a) and top view (b) respectively. The blue line is the estimated trajectory of the handheld device's antenna. The red dots are the actual positions of the unknown RFID-tags.



(a) x-axis and y-axis coordinates. (b) z-axis and x-axis coordinates.

Fig. 14. Experimental tag localization results. The red dots are the actual positions of the unknown RFID-tags. The black dots are the corresponding estimations, connected with green lines (the error) from their actual positions.

reader and a camera. The method is suitable for manual inventorying, where traditionally only the existence of an RFID tag is expected. The device calculates its own trajectory exploiting optical and RFID landmarks at known locations. Detection of the optical markers provide estimations of the positions of the reader, and phase measurements from the reference tags enable the calculation of the device's velocity. The position and velocity estimations are fused using a Kalman filter to get the final estimation of the devices 3D trajectory. As soon as the trajectory is estimated, the positions of all unknown tags in range are also estimated by solving a proper minimization problem. The proposed method is real-time.

Experimental results show that increased accuracy can be achieved by increasing the spatial density of the optical and RFID markers.

Future work will be focused on avoiding any installation of optical markers, but rather exploiting photos collected from the actual scene. Furthermore, as the accuracy of tags' localization greatly depends on the accuracy of the trajectory estimation, improved optical equipment is expected to increase the number of "quality" frames, thus improving the results. Different approaches to the pre-processing of the captured frames as well as how optical and RFID data are fused can be examined.

REFERENCES

[1] A. R. Chatzistefanou and A. G. Dimitriou, "Tag Localization by Handheld UHF RFID Reader and Optical Markers," 2022 IEEE 12th International Conference on RFID Technology and Applications (RFID-TA), 2022, pp. 9-12, doi: 10.1109/RFID-TA54958.2022.9924090.

[2] A. R. Chatzistefanou, A. Tzitzis, S. Megalou, G. Sergiadis and A. G. Dimitriou, "Target Localization by Mobile Handheld UHF RFID Reader and IMU," in IEEE Journal of Radio Frequency Identification, doi: 10.1109/JRFID.2022.3147539.

[3] E. DiGiampaolo and F. Martinelli, "A Robotic System for Localization of Passive UHF-RFID Tagged Objects on Shelves," in IEEE Sensors Journal, vol. 18, no. 20, pp. 8558-8568, 15 Oct.15, 2018, doi: 10.1109/JSEN.2018.2865339.

[4] E. DiGiampaolo and F. Martinelli, "Mobile Robot Localization Using the Phase of Passive UHF RFID Signals," in IEEE Transactions on

Fig. 15. The environment of the experiment presented in section III-C. The density of optical markers and reference tags was increased to improve the antenna trajectory estimation accuracy.

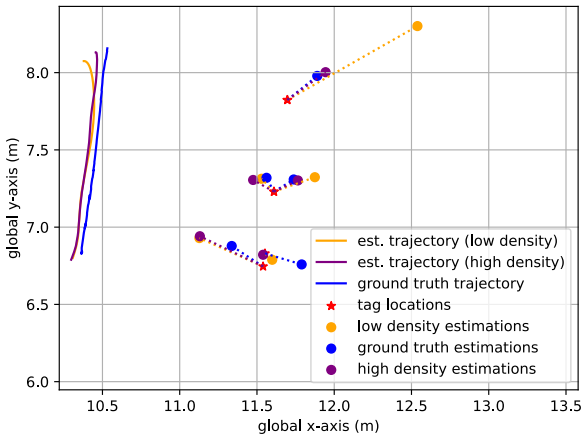


Fig. 16. Experimental tag localization results for different spatial densities of optical and RFID markers.

Industrial Electronics, vol. 61, no. 1, pp. 365-376, Jan. 2014, doi: 10.1109/TIE.2013.2248333.

[5] F. Shamsfakhr, A. Motroni, L. Palopoli, A. Buffi, P. Nepa, D. Fontanelli, "Robot localisation using UHF-RFID tags: A kalman smoother approach," in Sensors. 21. 717., 2021, 10.3390/s21030717.

[6] A. Motroni, A. Buffi, P. Nepa and B. Tellini, "Sensor-Fusion and Tracking Method for Indoor Vehicles With Low-Density UHF-RFID Tags," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-14, 2021, Art no. 8001314, doi: 10.1109/TIM.2020.3027926.

[7] B. Tao, H. Wu, Z. Gong, Z. Yin and H. Ding, "An RFID-Based Mobile Robot Localization Method Combining Phase Difference and Readability," in IEEE Transactions on Automation Science and Engineering, vol. 18, no. 3, pp. 1406-1416, July 2021, doi: 10.1109/TASE.2020.3006724.

[8] F. Martinelli, "A Robot Localization System Combining RSSI and Phase Shift in UHF-RFID Signals," in IEEE Transactions on Control Systems Technology, vol. 23, no. 5, pp. 1782-1796, Sept. 2015, doi: 10.1109/TCST.2014.2386777.

[9] H. Wu, B. Tao, Z. Gong, Z. Yin and H. Ding, "A Standalone RFID-Based Mobile Robot Navigation Method Using Single Passive Tag," in IEEE Transactions on Automation Science and Engineering, vol. 18, no. 4, pp. 1529-1537, Oct. 2021, doi: 10.1109/TASE.2020.3008187.

[10] D. Hahnel, W. Burgard, D. Fox and S. Thrun, "An efficient fast-SLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), 2003, pp. 206-211 vol.1, doi: 10.1109/IROS.2003.1250629.

[11] Filotheou, A., Tsardoulias, E., Dimitriou, A. et al. Pose Selection and Feedback Methods in Tandem Combinations of Particle Filters with Scan-Matching for 2D Mobile Robot Localisation. J Intell Robot Syst 100, 925–944 (2020). https://doi.org/10.1007/s10846-020-01253-6

[12] Filotheou, A., Tzitzis, A., Tsardoulias, E. et al. Passive Global Localisation of Mobile Robot via 2D Fourier-Mellin Invariant Matching. J Intell Robot Syst 104, 26 (2022). https://doi.org/10.1007/s10846-021-01535-7

[13] X. Zeng, G. He and Y. Zhuang, "B-Spline-Based Trajectory Estimation for Handheld LiDAR-SLAM Device," 2021 3rd International Conference on Industrial Artificial Intelligence (IAI), 2021, pp. 1-5, doi: 10.1109/IAI53119.2021.9619441.

[14] Z. Wang, M. Xu, N. Ye, F. Xiao, R. Wang and H. Huang, "Computer Vision-Assisted 3D Object Localization via COTS RFID Devices and a Monocular Camera," in IEEE Transactions on Mobile Computing, vol. 20, no. 3, pp. 893-908, 1 March 2021, doi: 10.1109/TMC.2019.2954830.

[15] J. Bayer and J. Faigl, "Handheld Localization Device for Indoor Environments," 2020 4th International Conference on Automation, Control and Robots (ICACR), 2020, pp. 60-64, doi: 10.1109/ICACR51161.2020.9265494.

[16] C. Duan, X. Rao, L. Yang and Y. Liu, "Fusing RFID and computer vision for fine-grained object tracking," IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, 2017, pp. 1-9, doi: 10.1109/INFOCOM.2017.8057161.

[17] A. Milella, D. Di Paola, G. Cicirelli and T. D'Orazio, "RFID tag bearing estimation for mobile robot localization," 2009 International Conference on Advanced Robotics, 2009, pp. 1-6.

[18] D. Fox, W. Burgard, and S. Thrun, "Markov Localization for Mobile Robots in Dynamic Environments," Journal of Artificial Intelligence Research, vol. 11, pp. 391-427, 1999.

[19] S. Kunkel, R. Bieber, Ming-Shih Huang and M. Vossiek, "A concept for infrastructure independent localization and augmented reality visualization of RFID tags," 2009 IEEE MTT-S International Microwave Workshop on Wireless Sensing, Local Positioning, and RFID, 2009, pp. 1-4, doi: 10.1109/IMWS2.2009.5307726.

[20] T. Kim, J. J. Song and J. Ri, "Localization of handheld UHF RFID reader using reference tags in 3D environment," 2014 IEEE International Wireless Symposium (IWS 2014), 2014, pp. 1-4, doi: 10.1109/IEEE-IWS.2014.6864210.

[21] A. Parr, R. Miesen, F. Kirsch and M. Vossiek, "A novel method for UHF RFID tag tracking based on acceleration data," 2012 IEEE International Conference on RFID (RFID), 2012, pp. 110-115, doi: 10.1109/RFID.2012.6193037.

[22] M. Miyazawa, Y. Igarashi and J. Mitsugi, "A feasibility study of plug-in type RF tag localization system for handheld UHF RFID reader/writers," 2015 21st Asia-Pacific Conference on Communications (APCC), 2015, pp. 680-685, doi: 10.1109/APCC.2015.7412594.

[23] A. Tzitzis et al., "Real-time 3D localization of RFID-tagged products by ground robots and drones with commercial off-the-shelf RFID equipment: Challenges and Solutions," 2020 IEEE International Conference on RFID (RFID), 2020, pp. 1-8, doi: 10.1109/RFID49298.2020.9244904.

[24] A. Tzitzis, A. Malama, V. Drakaki, A. Bletsas, T. Yioultsis and A. G. Dimitriou, "Real-Time, Robot-Based, 3D Localization of RFID Tags, by Transforming Phase Measurements to a Linear Optimization Problem," in IEEE Journal of Radio Frequency Identification, doi: 10.1109/JRFID.2021.3103393.

[25] P. Tripicchio et al., "A Synthetic Aperture UHF RFID Localization Method by Phase Unwrapping and Hyperbolic Intersection," in IEEE Transactions on Automation Science and Engineering, vol. 19, no. 2, pp. 933-945, April 2022, doi: 10.1109/TASE.2021.3057433.

[26] F. Bernardini et al., "Robot-Based Indoor Positioning of UHF-RFID Tags: The SAR Method With Multiple Trajectories," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-15, 2021, Art no. 8001415, doi: 10.1109/TIM.2020.3033728.

[27] J. Zhang, Y. Lyu, J. Patton, S. C. G. Periaswamy and T. Roppel, "BFVP: A Probabilistic UHF RFID Tag Localization Algorithm Using Bayesian Filter and a Variable Power RFID Model," in IEEE Transactions on Industrial Electronics, vol. 65, no. 10, pp. 8250-8259, Oct. 2018, doi: 10.1109/TIE.2018.2803720.

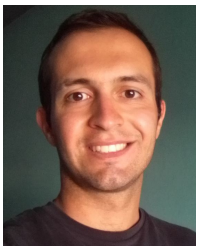
[28] L. Yang, Q. Liu, J. Xu, J. Hu and T. Song, "An Indoor RFID Location Algorithm Based on Support Vector Regression and Particle Swarm Optimization," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), 2018, pp. 1-6, doi: 10.1109/VTCFall.2018.8690994.

[29] S. J. Patel and M. J. Zawodniok, "3D Localization of RFID Antenna Tags Using Convolutional Neural Networks," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-11, 2022, Art no. 8001311, doi: 10.1109/TIM.2022.3146604.

[30] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. 2014. "Automatic generation and detection of highly reliable fiducial markers under occlusion". Pattern Recogn. 47, 6 (June 2014), 2280-2292. DOI=10.1016/j.patcog.2014.01.005

[31] G. Bradski. The OpenCV Library. Dr. Dobb's Journal of Software Tools, 2000.

[32] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. 2014. Tagoram: real-time tracking of mobile RFID tags to high precision using COTS devices. In Proceedings of the 20th annual international conference on Mobile computing and networking (MobiCom '14). Association for Computing Machinery, New York, NY, USA, 237–248. https://doi.org/10.1145/2639108.2639111



Aristidis Raptopoulos Chatzistefanou was born in Florina, Greece, in 1996. He received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki in 2019, where he is currently pursuing the Ph.D. degree, and working as a Research Assistant with the Aristotle University of Thessaloniki. His main research interests include RFID technology, localization and tracking techniques.

George Sergiadis received his diploma in Electrical Engineering from the Aristotle University of Thessaloniki, Greece, and his Ph.D. from "Ecole Nationale Supérieure des Télécommunications", in Paris, France.

He is currently full Professor at the Aristotle University of Thessaloniki, Greece, teaching Telecommunications and Biomedical Engineering, since 1985. He has been an August-Wilhelm Scheer visiting professor in TUM, Munich, Germany, a visiting researcher at Media Lab, MIT, also in IBMI, Munich, Germany, and a Distinguished Professor at SIBET, Suzhou, China, for several years.

His current research interests include signal and image processing, and medical imaging.

Antonis G. Dimitriou (S'01-M06-SM'14) received the diploma and the Ph.D. degree in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece, in 2001, and 2006 respectively. Since 2007, he is with the School of Electrical and Computer Engineering of AUTh, where he currently serves as a teaching and research faculty member.

He has participated in more than 20 research projects, 8 of which since 2007 as a principal investigator in the fields of Robotics, RFIDs, and Wireless Sensor Networks. He was the coordinator of project "RELIEF", involving continuous RFID inventorying through robotics and he is currently the coordinator of project "CULTUREID", where RFID equipment is installed inside the Archaeological Museum of Thessaloniki to monitor RFID tagged exhibits, track visitors and interact with a prototype RFID-enabled social robot. He was a Management Committee Member in the ICT COST Action IC301 "Wireless Power Transmission for Sustainable Electronics (WiPE)". He is the author or co-author of approximately 90 journal and conference papers.

Dr. Dimitriou was the recipient of the Ericsson Award of Excellence in Telecommunications in 2001, co-recipient of the student-paper award in the 2011 IEEE RFID-TA conference and finalist in the IEEE best paper award in the 2022 IEEE RFID conference. He received the "IEEE Wireless Communications Letters Exemplary Reviewer" award in 2012 and 2014. He is a Senior IEEE Member since February 2014. He is the co-chair of the IEEE Technical Committee on "Motion Capture and Localization" and Associate Editor in the IEEE Journal of Radio Frequency Identification. He also serves as a reviewer for major journals and as a TPC member for international conferences.