Digital Twin of Retail Stores with RFID Tags Localization

Junwei Ma[†], Xiangyu Wang[‡], Caleb Powell[‡], Jian Zhang[§], Shiwen Mao[†]
Senthilkumar CG Periaswamy[‡], and Justin Patton[‡]
[†]Dept. Electrical & Computer Engineering, Auburn University, Auburn, AL 36845-5201, USA
[‡]RFID Lab, Auburn University, Auburn, AL 36830, USA
[§]Department of Information Technology, Kennesaw State University, Marietta, GA 30144, USA
Email: {jzm0175, xzw0042, cmp0132}@auburn.edu, {jianzhang, smao}@ieee.org, {szc0089, jbp0033}@auburn.edu,

***Abstract*—In this work, we integrate digital twin technology with RFID localization to achieve real-time monitoring of physical items in a large-scale complex environment, such as warehouses and retail stores. To map the item-level realities into a digital environment, we proposed a sensor fusion technique that merges a 3D map created by RGB-D and tracking cameras with real-time RFID tag location estimation derived from our novel Bayesian filter approach. Unlike mainstream localization methods, which rely on phase or RSSI measurements, our proposed method leverages a fixed RF transmission power model. This approach extends localization capabilities to all existing RFID devices, offering a significant advancement over conventional techniques. As a result, the proposed method transforms any RFID device into a digital twin scanner with the support of RGB-D cameras. To evaluate the performance of the proposed method, we prototype the system with commercial off-the-shelf (COTS) equipment in two representative retail scenarios. The overall performance of the system is demonstrated in a mock retail apparel store covering an area of 207 m^2, while the quantitative experimental results are examined in a small-scale testbed to showcase the accuracy of item-level tag localization.**

***Index Terms*—Digital twin, radio-frequency identification (RFID), Bayesian filter, localization, Real-time appearance-based mapping (RTAB-Map)**

I. Introduction

With the rapid development of the Internet of Things (IoT), the *Radio-frequency identification* (RFID) technology has drawn growing attention from numerous fields [1]–[3], and has been utilized for a broad range of applications such as retailing, asset tracking, healthcare, and supply chain management [4]. In the last two decades, the RFID technology has been a popular choice for item localization due to its low cost and simplicity of deployment [5]. On the other hand, the concept of *digital twin* has recently been brought to the attention of RFID researchers considering its high potential for improving customer services in the retail industry [6]. A digital twin refers to a virtual representation (usually in the digital world) of a physical system, while the digital twinning technology aims at building a high-fidelity virtual model or depiction of a real-world entity or system, which involves both digital and physical elements [7]. With the integration of the RFID technology and digital twin, we can create precise models to better analyze the status of inventory and offer retailers an accurate and timely overview of the retail store for effective exploration and management.

With the rapid growth since 2016, the digital twin has been considered a critical component of many fields such as smart manufacturing. The innovative data streaming through digital twins could potentially benefit the real-world process [8]–[10]. Although the digital twin has shown great success in many areas of industry, the number of studies that have integrated RFID technology is limited, especially in relation to retail store and warehouse management. Some researchers transform a simulation model to establish a digital twin that could interact with incoming data and optimize the model efficiency. Braglia et al. [11] present an agent-based simulation model of a large paper warehouse, in which the forklifts and the pallets are identified by reading their attached Ultra High Frequency (UHF) RFID tags. The sensors update data for the locations of products and forklifts in the digital twin at certain time intervals. The simulation model (which is the digital twin) quickly develops new alternative scenarios, attempting to optimize the warehouse allocation and the routes that the forklifts should take to minimize the overall costs of warehouse management. Then, the optimized routes are shown to the drivers and are then carefully followed.

The amalgamation of robotics and digital twin technologies has been developed in the retail store and warehouse management scenarios. Various studies have been focused on deploying RFID techniques to localize items to benefit inventory tracking [12]. The prior works mainly adopted RFID tags as landmarks to accurately pinpoint the tagged products [13]. Maïzi and Bendavid developed an RFID-based digital twin system for inventory tracking, customer monitoring, and operations management [6]. Various machine learning techniques have been applied in operation optimization of digital twins. For instance, Pous et al. [14] designed a digital twin of a retail store that contains item information of product name and price. A ground RFID robot collected data for product inventory and localization, and then the data was fed into the 3D-like digital twin interface that allowed customers to browse virtually.

In this paper, we present a pilot model of retail store digital twin based on the Real-Time Appearance-Based Mapping (RTAB-map) [15] technology. Each product attached with an RFID tag was localized and marked in the proposed digital twin model. Our system deploys two RGB-D cameras to capture visual information and a handheld RFID reader to

scan RFID tags in the deployed environment. While the human operator is scanning RFID tags, both visual and localization data are fed into the digital twin model in the system. The 3D digital twin was constructed in real-time, and the localization results will be shown once the data collection is complete. Finally, all scanned tags are labeled in the 3D environment of the generated retail store digital twin. we prototype the system with commercial off-the-shelf (COTS) equipment in two representative retail scenarios. The overall performance of the system is demonstrated in a mock retail apparel store covering an area of 207 m2, while the quantitative experimental results are examined in a small-scale testbed to showcase the accuracy of item-level tag localization. The experiments illustrated accurate localization results, within 0.5 meters, were achieved even under strong multipath effects. With the integration of digital twin and RFID localization in such retail store environments, our proposed system could greatly benefit the efficiency of inventory and supply chain.

The rest of this paper is organized as follows. In Section II, we present the system design. In Section III, we present the experiment results. Section IV concludes this paper.

II. Methodology

In this section, the Bayesian filter is presented for passive UHF RFID tag localization. Then we introduce the fixed transmission power RF model for improved location estimation. Finally, we describe the RTAB-Map method and show how the digital twin is established to merge the map and localization results, including coordinate transformation from camera frames to global frames.

A. Bayesian filter-based RFID localization

Bayesian filtering is known as a classical method to estimate the position of RFID tags in a specific environment, which has also been widely used in robotics [12]. As a statistical approach, it is used to probabilistically determine the most likely position of RFID tags by accounting for multiple sources of uncertainty, including fluctuations in signal strength, multipath effects, and environmental impediments.

Assuming a noise-affected dynamic system, presenting a random variable of x_t at time point t. Bayesian filter aims to sequentially determine the uncertainty of a probability distribution over x_t, which is called belief $\mathcal{B}(x_t)$ that is based on a sequential observation z_t. The belief $\mathcal{B}(x_t)$ is determined by the posterior distribution state x_t and all the available observations $z_1, z_2, \cdots, z_t$ at time t, given by

$$\mathcal{B}(x_t) = p(x_t|z_1, z_2, \cdots, z_t), \tag{1}$$

In general, $\mathcal{B}(x_t)$ denotes the probability that the subject is in state x_t with all the known observations $z_1, z_2, \cdots, z_t$. The Bayesian filter can be divided into two stages in one update loop. The first stage is the prediction stage, in which the system gets a new observation at time t:

$$\overline{\mathcal{B}}(x_t) = \int P(x_t|x_{t-1})\mathcal{B}(x_{t-1})dx_{t-1}, \tag{2}$$

where $\mathcal{B}(x_{t-1})$ denotes the previous estimated probability at time $t-1$. $P(x_t|x_{t-1})$ is a motion model that indicates the probability of the subject moving from state x_{t-1} to state x_t. Since the new belief $\mathcal{B}(x_t)$ is generated, then the process comes to the second stage which is the update or correction:

$$\mathcal{B}(x_t) = \beta \cdot \overline{\mathcal{B}}(x_t) \cdot P(z_t|x_t), \tag{3}$$

where we have the observation model $P(z_t|x_t)$, which is the probability of the subject capturing observation z_t in state x_t.

In this study, it was assumed, without loss of generality, that the environment is static, which means that all the RFID tags are stationary and located in the detecting area during the experiments. As we discussed above, the user hold the RFID reader with two cameras mounted to collect data at various positions. Every read of RFID tag response was recorded and correlated to the specific pose in the predetermined global map. Then, after all the readings from the RFID tags are received, the locations will be estimated using the Bayesian filter. With the testing environment, we have state x as a three-dimensional (3D) vector indicating the location of an RFID tag. At each time t, we obtain observation z_t from the RFID reader, and $\overline{\mathcal{B}}(x_t) = \mathcal{B}(x_{t-1})$ based on the stationary assumption. Then we have the following equation:

$$\mathcal{B}(x_t) = \beta \cdot \mathcal{B}(x_{t-1}) \cdot P(z_t|x, r_t), \tag{4}$$

where r_t is the location of the RFID reader at time t. Based on (4), we have the observation model $P(z_t|x, r_t)$, it shows the probability of the RFID reader at location r_t, with the measurements of RFID response z_t from the tag's fixed location x. With the recursively updating algorithm, the Bayesian filter can determine the most likely location of the RFID tag.

A one-dimensional (1D) illustration of the Bayesian filter is shown in Fig. 1, and the detailed steps are presented in Algorithm 1. At the beginning of algorithm, based on all the observations from the RFID tag, the hypothetical location x can be estimated. Having the set of all observations $\{z_1, z_2, \cdots, z_t\}$ recorded for RFID tags, an individual observation z_t at time t indicates the probability of the hypothetical location in the entire sensing area. Considering a sensing area divided into equal-sized grids, the hypothetical location could be presented as the set of all the grids in the area. The posterior state could be accumulated as a union based on all the observations temporally. Finally, the estimated location of the RFID tag is determined according to the maximal belief, as

$$\hat{x} = \underset{x}{\operatorname{argmax}}\ \mathcal{B}(x). \tag{5}$$

Moreover, the geometrical center is adopted if there are multiple locations with the maximal belief $\hat{x}$.

Fig. 1 and Algorithm 1 show that the results of Bayesian filter mainly depend on the model $P(z_t|x, r_t)$. Considering the complex experimental environment, a simple RF signal radiation model may not be sufficient. Therefore, we apply an empirical model according to numerous observations during experiments. In general, the observations are captured by an

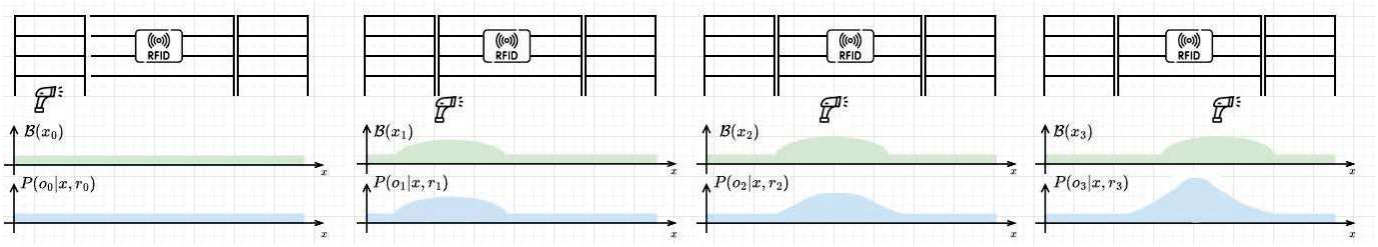

Fig. 1: An illustration of RFID tag localization with Bayesian filter. As the figure shows, the operator is holding an RFID reader to scan a tag at different locations. Each frame indicates the updated belief $\mathcal{B}(x_t)$ at time t.

Algorithm 1 Bayesian filter for RFID tag localization

Input: observations for an RFID tag and the reader pose ;
Output: the estimated tag location $\hat{x}$;

```
//M denotes the total number of observations
//N represents the total number of location grids in the entire sensing zone
Initialize the set of grids to represent hypothetical location x according to the observation set ;
x = {s_1 ∪ s_2 ··· ∪ s_M} ;
Hypothetical location grids s_t = {x_t^1, x_t^2, ..., x_t^N} based on observation z_t ;
for t = 1 : t do
    Update B(x_t) as in (4) ;
end for
//estimate the tag location
x̂ = argmax_x B(x) ;
return x̂ ;
```

RFID reader at a fixed position, while the RFID tag has dynamic locations with random orientation. The entire search zone is evenly gridded at a fixed spatial interval of 10 cm, and the tag is attached at the center of each grid during each recording. In each round of data collection, the RF transmit power is constant. By putting an RFID tag in each grid, we can observe multiple reads from the reader's antenna. The observation collection consists of 20 rounds of the experiment, and we establish a map of successful readings from each grid in the sensing zone. The probabilistic RFID observation model is acquired, in which a higher observation probability indicates a larger number of successful readings. More details on the observation experiment can be found in our previous work [12]. The Bayesian filter with a fixed RF transmit power model is developed based on the approaches in [16].

B. RTAB-Map based digital twin construction

Real-Time Appearance-Based Mapping (RTAB-Map) [15] is a powerful open-source approach of a Simultaneous Localization And Mapping (SLAM) solution [17]. It is widely used in robotics localization and mapping studies to provide real-time environment map construction based on sensors such as cameras and LiDAR.

In this work, we create digital twin models in small and large environments by deploying RTAB-Map with two cameras mounted on an RFID portable reader. Specifically, the RealSense D435 camera is capable of capturing both RGB and depth images (i.e., RGB-D), thus furnishing the essential data required for constructing a 3D map [18]. The RealSense T265, as a specialized tracking camera, can offer odometry data, enabling the localization and mapping of the entire scanning system [19]. Before data acquisition starts, two cameras ought to be properly calibrated. The procedure of calibration comprises the identification and adjustment of both intrinsic and extrinsic parameters of the cameras to establish a precise correspondence between pixel coordinates and the global coordinate [18]. Synchronization of the sensing process plays an essential role in ensuring the creation of an accurate digital twin. The objective is not only to achieve temporal alignment of RGB images, depth images, and odometry data by synchronizing the data streams from both cameras, but also to align the visual data from cameras with continuous RF signals. The visual features, including key points and descriptors, are extracted from the RGB-D images acquired with the D435 camera. The odometry data gathered from the t265 camera offers motion estimation that facilitate feature matching over consecutive frames and the identification of loop closures [17]. Loop closure detection denotes the crucial procedure of comparing the current visual data with previously acquired data. It enhances the digital twin's consistency and mitigates the drift over time when the system refines the map and the agent's estimated pose in the map. The RTAB-Map method leverages graph optimization to maximize the consistency and accuracy of the map based on the designed graph structure of inter-frame agent pose and landmarks.

To integrate the odometry data and the RGB-D data from both cameras, we can build a comprehensive 3D spatial map as a digital twin of the target region. However, the pose of two cameras mounted together on the RFID reader antenna is described in the t265's built-in frame coordinate [18]. In the data acquisition phase, the user carries the RFID reader and camera system in the 3D space with six degrees of freedom (6-DoF) [13]. Its pose is defined as the combination of position **T** and orientation **R**, which are three translation DoFs and three rotation DoFs, respectively [20]. We assume a predetermined frame coordinates as $\mathcal{L}$ and a pose **P** was assigned to this coordinate according to its relative position and orientation of

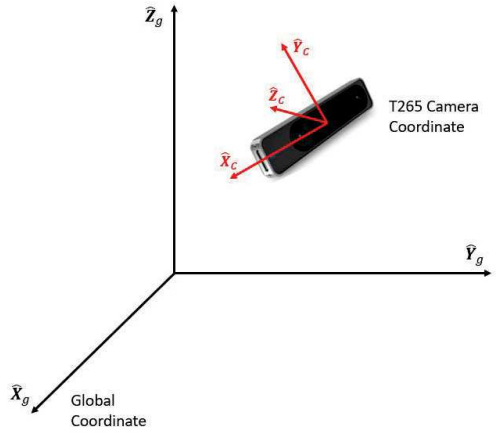

Fig. 2: The coordinate system: (i) the global coordinate and (ii) the camera local coordinate.

axes. A pose $\mathbf{P}$ related to the coordinate frame $\mathcal{L}$ is:

$$\mathbf{P} = [\mathbf{T}, \mathbf{R}]^T, \tag{6}$$

where position $\mathbf{T}$ is a 3D vector, orientation $\mathbf{R}$ is a 3D matrix, and $(\cdot)^T$ denotes the transpose operation.

As illustrated in Fig. 2, we defined two coordinate systems to represent the camera-reader pose in the 3D space. One is the camera's local coordinate c, which is the origin at the center point of the t265 camera. The other coordinate is the global coordinate g as the environment coordinate frame, with the origin located at a fixed location on the ground. The RFID reader detects a UHF tag at position $P_c = (x_c, y_c, z_c)^T$ in the local camera coordinate. To perform the rigid transformation from the camera coordinate to the global coordinate, the translation vector T_c^g and rotation matrix R_c^g are calculated. The 3D-translation represents the offset from the camera frame coordinate c relative to the global frame g, specified as :

$$T_c^g = [o_x^g - o_x^c, o_y^g - o_y^c, o_z^g - o_z^c]^T = [t_x, t_y, t_z]^T, \tag{7}$$

which $[o_x^c, o_y^c, o_z^c]$ and $[o_x^g, o_y^g, o_z^g]$ denote the origins of the camera coordinate frame and the global coordinate frame, respectively. The 3D-rotation matrix from the local camera coordinates to the global coordinates is specified as a 3×3 matrix, given by:

$$R_c^g = \begin{bmatrix} \hat{X}_c \cdot \hat{X}_g & \hat{Y}_c \cdot \hat{X}_g & \hat{Z}_c \cdot \hat{X}_g \\ \hat{X}_c \cdot \hat{Y}_g & \hat{Y}_c \cdot \hat{Y}_g & \hat{Z}_c \cdot \hat{Y}_g \\ \hat{X}_c \cdot \hat{Z}_g & \hat{Y}_c \cdot \hat{Z}_g & \hat{Z}_c \cdot \hat{Z}_g \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix},$$

where $\hat{X}_c$, $\hat{Y}_c$, and $\hat{Z}_c$ denote the unit vector of local camera coordinate axes and $\hat{X}_g$, $\hat{Y}_g$, $\hat{Z}_g$ denote the unit vector of global coordinate axes. Based on the derived translation vector T_c^g and rotation matrix R_c^g, the frame coordinate transformation from local camera to global is given by:

$$\begin{aligned} P_g &= R_c^g \cdot P_c + T_c^g \\ &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} P_c + [t_x, t_y, t_z]^T, \end{aligned} \tag{8}$$

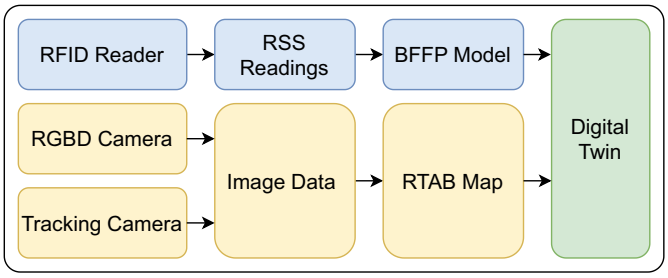

Fig. 3: System architecture of proposed approach: blue blocks represent the RFID pipeline, whereas the yellow blocks are the visual components.

By exporting the transformed tag locations to the RTAB-Map built 3D real-time map, we established the digital twin of the target area, which contains the visual features, structural geometry, and the RFID tags localization results. The system architecture of the proposed method is illustrated in Fig. 3.

III. Experiments and Results

A. Experimental setup

To evaluate the constructed digital twin, we perform a series of experiments at the RFID Lab of Auburn University. The experiment environment is the mock retail apparel store within an enclosed area of about 12×17 m^2, which has 10 shelves of products and furniture including tables and sofas. The overall layout of the retail store area are shown in Fig. 5 and a photo is shown in Fig. 6(a). There are 427 items with UHF RFID tags attached, including 33 pairs of shoes, 18 dresses, 22 pairs of short pants, 26 khakis, 124 T-shirts, and 204 jeans. As the top view of the layout map in Fig. 5 shows, the shoes are placed on a wooden shelf that is marked by block number 1. The dresses are distributed on two metal racks of numbers 8 and 9. Number 6 is a metal rack on which most of the T-shirts are hung. Some short pants and T-shirts are distributed on the metal racks, which are denoted as numbers 5 and 7, respectively. The block marked as number 10 is a metal shelf hung on the wall near the back door.

As shown in Fig. 6(b), a Realsense D435 RGB-D camera and a Realsense T265 tracking camera are mounted on top of the the Zebra RFD8500 Bluetooth Handheld UHF RFID Reader as our prototype of the RF-visual sensing component. The RGB-D camera and the tracking camera are connected to a host computer through USB3 Type-C and USB Micro B cables respectively. Connection between the Zebra RFD8500 reader [21] and the host computer is established via Bluetooth 2.1. ROS Noetic is leveraged to synchronize all the sensors in the proposed prototype.

The passive UHF RFID tag we used is the Avery Dennison AD-237 [22], which is widely utilized in retail inventory. It has an IC-type Impinj Monza R6 and 96-bit EPC for identification. In the large-scale experiment, all the merchandise item is attached with an AD-237 tag that is randomly orientated. The tags are allocated at a height from $1 \sim 1.5$ meters above the floor of the retail store. The system is not restricted to

Fig. 4: The overview of all the localized RFID tags in the 3D map.

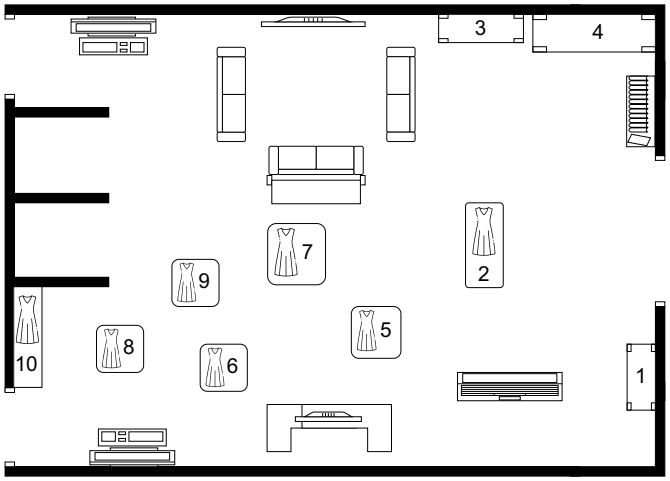

Fig. 5: The overall layout of the mock retail store. The indexed blocks indicate the locations of tagged items.

(a) (b)

Fig. 6: (a) The retail store environment. (b) The RF-visual sensing component: Realsense D435 RGB-D camera (Top), Realsense T265 tracking camera (Middle) and the Zebra RFD8500 Bluetooth Handheld UHF RFID Reader (Bottom).

any specific tag model and other types of tags have been recognized during the data acquisition process.

To evaluate the performance of the tag localization in the digital twin, we also conduct a small-scale experiment in the same environment to obtain quantitative results. In this scenario, the shoe shelf is scanned and its digital twin is created. The actual location of the tags attached to the shoes are carefully measured to serve as the ground truth, facilitating the performance evaluation of the proposed system.

B. Experimental results and discussions

Fig. 4 shows an example of the digital twin we built for the retail store to evaluate the performance of the proposed prototype in a large-scale scenario. In Fig. 4, the 3D map involves the reception desk, four racks (numbers 5-9), and block number 10, which is a shelf of jeans. All the RFID tags are localized and shown in the figure as red dots. This digital twin was created by the proposed method performed with the equipment shown in Fig. 6. In the data collection phase, users carry the equipment along with a host laptop computer to explore the target area. RGB-D images and the RFID tag readings are captured simultaneously to build the real-time digital twin of the retail store. All the racks are scanned in 360 degrees at various heights. The entire process of the experiment is accomplished within 5 minutes. The large-scale scenario experiment was performed in a heavy clutter. Since it is an application of manual inventory counts, our RF model does not consider the situation of non-line-of-sight.

(a) (b)

Fig. 7: (a) The digital twin of a shoe shelf. (b) The shoe shelf.

Furthermore, Fig. 7 illustrates the generated digital twin for a small-scale scenario. Similar to the digital twin of the retail store, all the tags are localized and marked as red spheres with the last four digits of individual EPC. To evaluate the performance of the item-level tag localization, tag locations are manually measured to establish the ground truth. According to Fig. 7, thirty three RFID tags are deployed in this experiment and are all detected and localized. Furthermore, all the tags are distributed within a reasonable range. The average localization error is 0.304 m and the maximum error is 0.600 m in this experiment.

IV. Conclusions

In this work, we developed a portable system to establish a prototype model of the digital twin of a retail store. Two cameras and a portable RFID reader were utilized to collect data. A Bayesian filter-based RFID tag localization algorithm was adopted in the proposed system. As a result, the digital twin model of a retail store can be easily created. The performance of tag localization was evaluated in the smalls-scale experiment, with an average error of 0.304 m. The digital twin could improve customers' experience at the store as well as the efficiency of inventory management.

Acknowledgment

This work is supported in part by the NSF under Grants CCSS-2245608 and CCSS-2245607.

References

[1] J. Zhang, S. Mao, S. Periaswamy, and J. Patton, "Standards for passive UHF RFID," *ACM GetMobile*, vol. 23, no. 3, pp. 10–15, Sept. 2019.

[2] X. Wang, J. Zhang, Z. Yu, S. Mao, S. C. Periaswamy, and J. Patton, "On remote temperature sensing using commercial uhf rfid tags," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 715–10 727, Sept. 2019.

[3] X. Wang, J. Zhang, S. Mao, S. Periaswamy, and J. Patton, "MulTLoc: RF hologram tensor filteing and upscaling for indoor localization using multiple UHF passive RFID tags," in *Proc. ICCCN 2021*, Athens, Greece, July 2021, pp. 1–9.

[4] E. Elbasani, P. Siriporn, and J. S. Choi, "A survey on RFID in Industry 4.0," in *Internet of Things for Industry 4.0*, G. Kanagachidambaresan, R. Anand, E. Balasubramanian, and V. Mahima, Eds. Cham, Switzerland: Springer, 2020, ch. 1, pp. 1–16.

[5] B. C. Hardgrave, J. A. Aloysius, and S. Goyal, "Rfid-enabled visibility and retail inventory record inaccuracy: Experiments in the field," *Production and Operations Management*, vol. 22, no. 4, pp. 843–856, Mar. 2013.

[6] Y. Maizi and Y. Bendavid, "Building a digital twin for iot smart stores: A case in retail and apparel industry," *International Journal of Simulation and Process Modelling*, vol. 16, no. 2, pp. 147–160, June 2021.

[7] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP journal of manufacturing science and technology*, vol. 29, pp. 36–52, May 2020.

[8] F. Tao, H. Zhang, A. Liu, and A. Y. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, Sept. 2018.

[9] G. Moiceanu and G. Paraschiv, "Digital twin and smart manufacturing in industries: A bibliometric analysis with a focus on Industry 4.0," *MDPI Sensors*, vol. 22, no. 4, p. 1388, Feb. 2022.

[10] F. Quazi, A. Vaequez-Castro, D. Niyato, S. Mao, and Z. Han, "Building delay-tolerant digital twins for cislunar operations using age of synchronization," in *Proc. IEEE ICC 2024 Workshops*, Denver, CO, June 2024, pp. 1–6.

[11] M. Braglia, R. Gabbrielli, M. Frosolini, L. Marrazzini, and L. Padellini, "Using RFID technology and discrete-events, agent-based simulation tools to build digital-twins of large warehouses," in *Proc. 2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, Pisa, Italy, Sept. 2019, pp. 464–469.

[12] J. Zhang, Y. Lyu, J. Patton, S. C. Periaswamy, and T. Roppel, "BFVP: A probabilistic UHF RFID tag localization algorithm using Bayesian filter and a variable power RFID model," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, pp. 8250–8259, Feb. 2018.

[13] J. Zhang, X. Wang, Z. Yu, Y. Lyu, S. Mao, S. C. Periaswamy, J. Patton, and X. Wang, "Robust rfid based 6-dof localization for unmanned aerial vehicles," *IEEE Access Journal*, vol. 7, pp. 77 348–77 361, June 2019.

[14] R. Pous, M. Chindemi, and A. Alajami, "Showing products in a retail store digital twin with item location captured by an rfid robot," in *Proc. 13th IEEE International Conference on RFID Technology and Applications (RFID-TA)*, Aveiro, Portugal, Sept. 2023, pp. 189–192.

[15] M. Labbe and F. Michaud, "Appearance-based loop closure detection for online large-scale and long-term operation," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 734–745, Feb. 2013.

[16] I. Ehrenberg, C. Floerkemeier, and S. Sarma, "Inventory management with an rfid-equipped mobile robot," in *Proc. 2007 IEEE International Conference on Automation Science and Engineering*, Scottsdale, AZ, Oct. 2007, pp. 1020–1026.

[17] M. Labbé and F. Michaud, "RTAB-map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *Journal of Field Robotics*, vol. 36, no. 2, pp. 416–446, Oct. 2019.

[18] L. Keselman, J. Iselin Woodfill, A. Grunnet-Jepsen, and A. Bhowmik, "Intel realsense stereoscopic depth cameras," in *Proc. IEEE CVPR 2017 Workshops*, Honolulu, HI, July 2017, pp. 1–10.

[19] P. Schmidt, J. Scaife, M. Harville, S. Liman, and A. Ahmed, "Intel® realsense™ tracking camera t265 and intel® realsense™ depth camera d435-tracking and depth," accessed: Mar. 2024. [Online]. Available: https://dev.intelrealsense.com/docs/depth-and-tracking-cameras-alignment

[20] J. Zhang, Z. Yu, X. Wang, Y. Lyu, S. Mao, S. C. Periaswamy, J. Patton, and X. Wang, "RFHUI: An RFID based human-unmanned aerial vehicle interaction system in an indoor environment," *Elsevier/KeAi Digital Communications and Networks*, vol. 6, no. 1, pp. 14–22, May 2020.

[21] Z. Technologies, "Zebra RFD8500 Handheld RFID," accessed: Mar. 2024. [Online]. Available: https://www.zebra.com/us/en/products/rfid/rfid-handhelds/rfd8500.html

[22] A. Dennison, "Avery Dennison Smartrac AD-237r6 UHF RFID," accessed: Mar. 2024. [Online]. Available: https://www.atlasrfidstore.com/avery-dennison-smartrac-ad-237r6-uhf-rfid-white-wet-inlay-monza-r6/