Exploiting Synergies between Augmented Reality and RFIDs for Item Localization and Retrieval

Tara Boroushaki[1], Maisy Lam[1], Weitung Chen[1], Laura Dodds[1], Aline Eid[2], Fadel Adib[1,3]
[1] *Massachusetts Institute of Technology,* [2] *University of Michigan,* [3] *Cartesian Systems*

***Abstract*—Locating RFID-tagged items in the environment and guiding humans to retrieve the tagged items is an important problem in the RFID community. This paper explores how to exploit synergies between Augmented Reality (AR) headsets and RFID localization to help solve this problem by improving both user experience and localization accuracy. Using fundamental mathematical formulations for RFID localization, we derive confidence metrics and display guidance to the user to improve their experience and enable them to retrieve items faster. We build our primitives into an end-to-end system, RF-AR, and show that it achieves 8.6 cm median localization accuracy within 76 seconds and enables 55% faster retrieval than state-of-the-art past systems. Our results demonstrate that AR-based "human-in-the-loop" designs can make the localization task more accurate and efficient, and thus holds the potential to improve processes where items need to be retrieved quickly, such as in manufacturing, retail, and warehousing.**

***Index Terms*—Augmented Reality, Virtual Reality, RFID Localization, User Interface, Human-in-the-loop, RF sensing**

I. Introduction

Augmented Reality (AR) and Virtual Reality (VR) enhance the interactions between humans and computers by offering an effective user interface. For example, AR glasses can significantly streamline operations in logistics and warehousing by visualizing the incoming customer orders and guiding the operators in the environment. In manufacturing, AR/VR glasses can display the next steps and tools to workers to improve accuracy and efficiency. Because of this immense potential to revolutionize the human-machine interface and the potential applications in future industries, substantial investments have been made in developing AR/VR headsets by governments and big tech companies.

Current AR/VR headsets are equipped with various sensors such as RGB and depth cameras, inertial sensors, and microphone arrays. While these devices are capable of hand tracking, mapping, and self-localization in their environment, their perception is as limited as humans. For instance, in the case of a store associate attempting to locate and retrieve a specific package from a customer's order, the AR/VR headset is unable to locate the package, guide the wearer towards it, or assist with the task unless there is a clear line of sight from the headset cameras to the packages barcodes.

One approach to enable new capabilities for AR/VR headsets is to leverage Radio Frequency (RF) signals. Since RF signals can traverse boxes and walls, they extend the perception beyond the line of sight. More specifically, RF sensing can be utilized to identify and sense Radio Frequency Identification (RFID) tagged items in the environment. Passive RFID tags are around 3-5 cents and are widely used.

Past work exploring the combination of RF and AR technologies have designed the wireless localization and AR

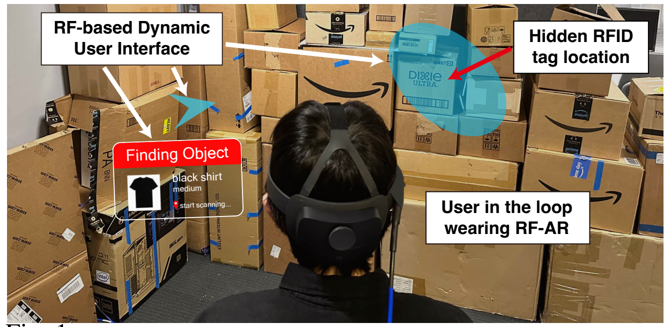


Fig. 1: **RF-AR.** RF-AR incorporates a dynamic RF-based user interface and brings the user in the loop to locate RFID tagged items. It uses color coding to communicate signal strength, projects an elliptical hologram to show the region on interest for the tag, and displays an arrow to help the user optimize their path to localize the target item.

system as separate entities. RF-AR, however, considers how to synergistally combine the two to deliver new functionalities to improve the RFID localization accuracy as well as the user experience.

In this paper, we present RF-AR, an AR system that seamlessly integrates RF perception and localization into the user experience. Our system introduces a new "human-in-the-loop" design and provides a dynamic RF-based user interface to help users efficiently and accurately localize hidden RFID tags in the environment. Particularly, RF measurements from different locations are required to locate target RFID tags, and it is especially critical to achieve proper measurement aperture for accurate and quick tag localization [1]–[3]. However, it is quite complicated for a typical AR user to optimize their path or measurement aperture without prior knowledge of the target item's location and a complete understanding of RF measurements and localization techniques. Thus, the key challenge is integrating RF perception into a user-centered design such that we abstract away all the technical details but still provide meaningful information that can be understood by a user with no RF knowledge.

To overcome these challenges, RF-AR introduces two main innovations:

- **Path Optimization:** RF-AR predicts the quality of RFID measurements in potential trajectories, determines the optimal one, and directs the user along that path, as shown in Fig. 1. RF-AR creates a candidate list of next possible locations for a user to walk to. For every candidate location, RF-AR considers: 1) the reduction in Dilution of Precision (DoP) [4] that a measurement taken at that location would yield, 2) the distance to the estimated target item location, and 3) the expected Signal-to-Noise Ratio (SNR).
- **Dynamic RF-based User Interface:** RF-AR introduces a dynamic RF-based user interface that provides essential RF

information while abstracting away the complicated technical details. RF-AR employs simple visualizations to provide seamless feedback to the user in a way that biases their trajectories toward the tag location, and brings them in the loop for localizing RFID tagged objects as shown in Fig. 1.

RF-AR explores multiple user interface implementations of varying levels of complexity to determine whether more sophisticated and involved displays better assist the user. To do this, RF-AR tests combinations of different prompts:

1) Color coding to communicate to the user whether they are successfully receiving RF measurement with an acceptable SNR
2) Visualizing a hologram to picture the region of confidence around the estimated location of the target item
3) Guiding the user with an arrow to follow the optimized trajectory

We implemented an end-to-end prototype of RF-AR, tested our system on 20 users[1], and evaluated its performance in 80 trials. Using our system, users were able to locate the hidden RFID tagged target items with 8.6 cm accuracy within a median of 76 sec. We also compared our performance to a baseline [2], showing that with the baseline, it takes users 55% longer (118 sec) to locate the RFID tag compared to RF-AR.

II. Related Work

Prior work that fuses RF and AR have designed RF sensing modules as separate entities and either completely keep the user out of the loop or fail to create a sophisticated system to optimize their task. For example, existing RFID-AR systems leverage their "AR" smart phones or separate monitors for the sole purpose of visualizing primitive displays of tagged items in the environment [5]–[9]. Other work providing information to users via glasses or headsets for non RFID localization tasks such as assembly, picking or walking either provides no visual display to the user [10], needs pre-prepared infrastructure [11], or cannot actively guide the user but rather confirm correct action following task completion [7].

More, recent work has explored enabling RFID sensing on an AR headset by leveraging a custom designed antenna and natural human motion to localize RFID tags [2]. Practically, this system, called X-AR, has minimal interaction with the user and reduces them to mere headset carriers, which fails to recognize the purpose of AR devices to incorporate users as active participants within the system, rather than passive carriers of the device, similar to a hand-held RFID reader.

III. Primer: Wideband SAR Localization

Synthetic Aperture Radar (SAR) is a technique for localization and imaging. In contrast to standard antenna arrays, SAR relies on a single antenna moved to multiple locations to collect measurements and emulate an antenna array. To carry out SAR, RF-AR leverages an antenna mounted on the AR headset and captures measurements as the user walks throughout the environment. The location of the antenna is estimated by the visual-inertial odometry (VIO) system of the AR headset. When a passive RFID tag powers up and responds with its identifiers, RF-AR uses this response to estimate the wireless channel $h(t)$ as $h = \sum_t y(t)x^*(t)$ where $x^*(t)$ is the conjugate of the transmitted signal $x(t)$, and $y(t)$ is the received signal.

We can then estimate the power P received from every point in space based on the estimated wireless channel using the following equation. It is worth noting that RF-AR exploits frequency diversity by taking measurements over a wideband of frequencies to improve the localization accuracy. Formally:

$$P(x,y,z) = \left\| \frac{1}{K}\frac{1}{N}\sum_{j=1}^{K}\sum_{i=1}^{N} h_{i,j} e^{\frac{2\pi d_i(x,y,z)}{\lambda_j}} \right\| \tag{1}$$

where N is the total number of measurements, K is the total number of frequencies, $h_{i,j}$ is the channel estimation of the i^{th} location with the j^{th} frequency, d_i is the round trip distance from (x,y,z) to the i^{th} location, and λ_j is the wavelength of j^{th} frequency. Finally, the location of the tag, p_{tag} can be determined using the following formula:

$$p_{tag} = \texttt{argmax}_{(x,y,z)}(P(x,y,z)) \tag{2}$$

IV. Path Optimization

In this section, we describe RF-AR's approach for path optimization. Remember that RF-AR's goal is to optimize RFID localization by bringing the user into the loop. As a first step, RF-AR has to determine what would be the optimal trajectory for the user. To do this, every time RF-AR receives a new RF measurement, it creates a list possible next locations for a user to walk to. It predicts the quality of an RF measurements at each candidate location, and then directs the user towards the best location with a holographic arrow. To estimate the quality of an RF measurement for target item localization, RF-AR analyzes multiple factors: 1) DoP, 2) SNR, and 3) distance to the estimated target location. In this section, we elaborate on each factor separately.

A. Dilution of Precision

To improve the accuracy of RFID tag localization, RF measurements should be taken over a wide aperture. This is because each RF measurement inherently suffers from some level of error. When combining multiple RF measurements as explained in §III, these errors can accumulate and significantly increase the localization error. This phenomenon is called DoP [4]. When RF measurements are taken over a wider aperture, the DoP will be smaller, meaning that the effect of small errors in each RF measurement on the final localization accuracy will be less significant.

Formally, we can calculate the DoP value based on RF measurement positions as follows:

$$A = \begin{bmatrix} \frac{x_1-x_p}{R_1} & \frac{y_1-y_p}{R_1} & \frac{z_1-z_p}{R_1} \\ \frac{x_2-x_p}{R_2} & \frac{y_2-y_p}{R_2} & \frac{z_2-z_p}{R_2} \\ \vdots & \vdots & \vdots \\ \frac{x_n-x_p}{R_n} & \frac{y_n-y_p}{R_n} & \frac{z_n-z_p}{R_n} \end{bmatrix} \tag{3}$$

[1]This study was approved by the institution's IRB.

$$Q = \left(A^T A\right)^{-1}, \quad DOP = \sqrt{tr(Q)} \tag{4}$$

where (x_i, y_i, z_i) correspond to the location of the antenna at time i, (x_p, y_p, z_p) is the estimated target RFID tag location calculated from the collected RF measurements described in §III, R_i is the distance from (x_i, y_i, z_i) to (x_p, y_p, z_p), and $tr(.)$ is the trace of matrix.

RF-AR uses the DoP as a metric to evaluate the potential of a measurement taken at a given position to aid in localizing the tag. First, to determine a position for consideration, RF-AR uses information regarding the speed the user is walking at to estimate the location of the user at time t. For example, if the user is at a current location $\vec{p_c} = (x_c, y_c, z_c)$ while moving at the speed v in direction θ_c, the future measurement position $\vec{p_f}$ after a period time t_a can be estimated as:

$$\vec{r}(\theta_c, v, t_a) = [vt_a \cos(\theta_c), vt_a \sin(\theta_c), 0] \tag{5}$$

$$\vec{p_f}(\theta_c, v, t_a) = \vec{p_c} + \vec{r}(\theta_c, v, t_a) \tag{6}$$

At the position $\vec{p_c}$, RF-AR calculates how much the DoP would change if the user moves to $\vec{p_f}$. The improvement in DoP can be described by the function $C_1(\theta_c, v, t_a)$ defined as:

$$C_1(\theta_c, v, t_a) = DOP_{\{\vec{p_1}, \vec{p_2}, ..., \vec{p_c}, \vec{p_f}\}} - DOP_{\{\vec{p_1}, \vec{p_2}, ..., \vec{p_c}\}} \tag{7}$$

where $\{\vec{p_1}, \vec{p_2}, \dots, \vec{p_c}\}$ are past measurement positions and p_f is the candidate future measurement position.

B. SNR

Another important metric for evaluating the quality of RF measurements is SNR. Without sufficient SNR, an RF measurement is not able to accurately estimate the wireless channel and help determine the location of the RFID tag as detailed in Sec. §III. A measurements' SNR can be influenced by multiple factors, including the distance and angular separation between the antenna and the target RFID tag, the antenna's radiation pattern and gain, and environmental noise and interference. Since the environmental multipath and noise is hard to predict and the antenna gain does not change, in this subsection, we focus on estimating the quality of SNR at a candidate position by considering the angular separation between the user and the estimated location of the RFID tag. It is important to note that we assume a user is looking in the direction they are walking.

Recall that RF-AR has an antenna on its headset's visor. As shown in Fig. 2a, if an RFID tag is in front of the headset and within the field of view of antenna, the RF measurement should have good SNR. However, as shown in Fig. 2b, if an RFID tag is outside the field of view of the antenna, RF-AR cannot obtain an RF measurement with acceptable SNRs. RF-AR takes this understanding into account when estimating the quality of future RF measurements.

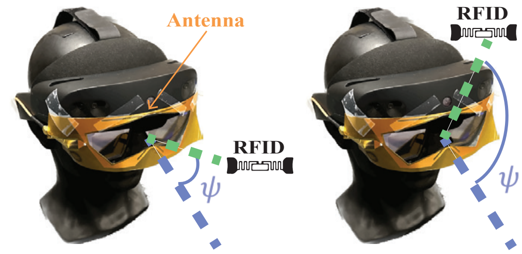
First, RF-AR calculates the angle ψ between the direction normal to the headset visor and its estimate of the target RFID location as:

$$\vec{r_h}(\theta_c) = [\cos(\theta_c), \sin(\theta_c), 0] \tag{8}$$

$$\vec{r_{tg}} = \vec{p_{tg}} - \vec{p_c} \tag{9}$$

$$\psi(\theta_c) = \cos^{-1} \frac{\vec{r_{tg}} \cdot \vec{r_h}(\theta_c)}{\|\vec{r_{tg}}\| \|\vec{r_h}(\theta_c)\|} \tag{10}$$

where $\vec{r_h}$ is the heading vector of the user walking in direction θ_c, and $\vec{r_{tg}}$ is the vector from the current user location to the estimated location. Based on ψ, RF-AR estimates the quality of SNR and scores it as follows:



(a) High SNR Estimate (b) Low SNR Estimate

Fig. 2: **Varied RFID Tag and AR headset Configurations.**

$$C_2(\theta_c) = |\sin(\psi(\theta_c))|, \quad \psi \in [-\frac{\pi}{2}, \frac{\pi}{2}] \tag{11}$$

where C_2 is the estimate quality of RF measurement SNR.

C. Distance to the target

An important factor that impacts the efficiency of finding hidden objects is the distance the users have to walk. Thus, RF-AR evaluates the quality of future RF measurement locations by also considering their distance to the target location. Specifically, the distance component of measurement quality can be formulated as:

$$C_3(\theta_c, v, t_a) = ||\vec{p_{tg}} - \vec{p_f}|| \tag{12}$$

where $\vec{p_{tg}}$ is the RF-AR's estimated location of the target RFID tag and $\vec{p_f}$ is the future measurement position as described in Eq.6.

D. Total Cost function

Ultimately, RF-AR incorporates all the factors mentioned above to decide on optimal candidate location to guide the user to. In practice, RF-AR samples θ_c from a discrete set of angles S_θ, and calculates the candidate RF measurement locations usiing to Eq. 6. To find the optimal guided direction θ_c^*, RF-AR minimizes the following cost function:

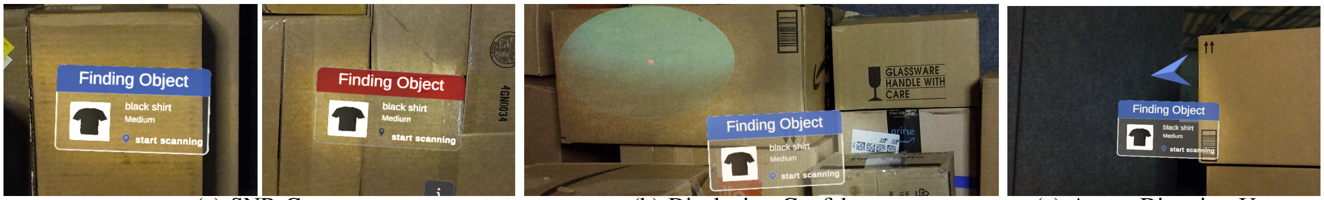
$$\theta_c^* = \arg\min_{\theta_c} C_1(\theta_c, v, t_a) + C_2(\theta_c) + C_3(\theta_c, v, t_a) \tag{13}$$

Based on the found θ_c^*, RF-AR then guides users towards the optimal path using a holographic arrow as we explain in the following section.

V. RF-Based Dynamic User Interface

In the previous sections, we summarized the SAR techniques for localization and described our path optimization algorithm. In this section, we describe how our system translates these concepts and creates an interactive and dynamic user display for optimized RFID localization.

Our goal is to create visual cues and holographic prompts that abstract away all the technical detail of RF measurements while still enabling the user to efficiently find hidden RFID tagged objects. Specifically, our focus is on providing the user with actionable prompts that reduce the overall time and distance traveled needed for RFID localization. Our design explores various RF-based user interfaces of varying degrees of sophistication. We describe the components of the user interfaces below.



(a) SNR Cue (b) Displaying Confidence (c) Arrow Directing Users

Fig. 3: **RF-based Dynamic User Interface.** (a) Examples of SNR visual cues given to the users. The system displays a blue color prompt to the user when the SNR is greater than μ, and changes the color to red when the SNR is smaller than μ. (b) A holographic ellipsoid communicates the current estimated location and its confidence to the user (c) An arrow that shows the direction to the next optimal location suggested by RF-AR's path optimization algorithm.

A. SNR Cue

Our system's first visual cue exploits the SNR to create an indicator for the user. As previously mentioned, low SNR prevents proper estimation of the wireless channel. Recall from §III that for each measurement, RF-AR performs multiple channel estimations over 200 MHz of bandwidth and then averages the result. Thus, we assume that a low average SNR at a position is not the result of environmental interference (which is unlikely to manifest at all frequencies), but rather the result of the user (and the visor-mounted antenna) facing a direction where the tag is not located.

By communicating the SNR status through a color-coded floating widget, as in Figure 3a, users can rule out possible tag locations and position themselves to optimize channel measurements. If the SNR is larger than μ, the system shows the user blue color, whereas if the SNR is smaller than μ, it shows the user red color for warning. We select μ such that RF measurements above the threshold[2] provide useful channel information. This simple and intuitive system enables users to quickly and easily identify areas with optimal RFID signal quality, allowing for efficient and accurate tag localization.

B. Displaying Confidence

In our AR application, we also visualize the concept of confidence in the estimated location of the RFID. Recall that SAR computes the power at each location in space based on Eq 1. To quantify the confidence of SAR in correctly determining the peak power location (corresponding to the target RFID), we select (x, y, z) points in space where the calculated power falls within 0.75dB of the peak power. We then extract the maximum distance along the x,y,z dimension for the selected points. When the area of these points is very large, it shows that SAR has not been able to narrow down the location of the RFID tag with reasonable confidence. Since these points tend to cluster into a 3D ellipse shape around the peak power, we display to the user a holographic ellipsoid fit to the extracted x,y,z axis, which is shown as the transparent blue ellipsoid in Fig. 3b. We model the ellipsoid in the AR display using the standard equations for Cartesian coordinate systems:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{14}$$

where a, b and c are the lengths of the semi-axis of the ellipsoid that correspond to the extracted axis dimensions.

As the confidence interval size updates with the addition of new measurements, the holograph of the ellipsoid updates as well. The RFID tag is considered confidently localized when the axis dimensions of the ellipsoid fall below (τ_x, τ_y, τ_z) [3], corresponding to a strong confidence in the predicted location similar to past work [2].

C. Directing Users to Optimal Path

Section §IV describes how we select the optimal next location for a measurement to increase the accuracy and efficiency of RFID localization. In order to direct users to this location, we display a blue arrow right above the floating canvas, as seen in Fig. 3c. This arrow updates every frame to ensure that it is always pointing the user in the correct direction as the user moves and walks through the space. The arrow also updates its pointing direction when new directions from path optimization are sent. The pointing vector of the arrow is calculated every frame as $\vec{q} =< x_n - x_h, y_n - y_h, z_n - z_h >$ where (x_h, y_h, z_h) are the coordinates just above the holographic floating canvas (base of the arrow) and (x_n, y_n, z_n) is the location that the system guides the user to, which is formulated as $\vec{p_n}(\theta_c^*, v, t_a)$ and is defined in Eq. 6. To make sure that the arrow does not point in an angled upward or downward direction that would confuse users, we set z_h equal to the height of the user's head, which is estimated by the VIO self-tracking of the AR device.

VI. Implementation

We programmed a custom application using Unity3D to display all user interfaces detailed in §V. Front end application graphics were designed in Figma and Adobe and then imported to Unity3D. Scripts to support application functionality were written in C# in Visual Studio IDE. We deployed our application on a Microsoft Hololens 2. We used the same antenna and hardware setup as [2] using BladeRFs, Raspberry Pi, and power splitters. We tested our device using standard off-the-shelf UHF RFID tags placed in cardboard boxes at different locations. We implemented all processing outlined in §III in C++ and python on the edge server which is an Ubuntu 20.04 machine with an Intel(R) Core(TM) i9-10900X CPU @ 3.70GHz. We implement code in Python on the Rasberry Pi to stream RFID channel measurements from BladeRFs to the edge server. The AR headset's UI then updates according to the received messages from the edge server.

VII. Evaluation

We evaluated RF-AR in an indoor environment that mimics a warehouse. Figure 1 shows our evaluation environment, which includes a number of stacked boxes. We recruited 20 participants (14 males and 6 females, aged 22-34 years old)

[2]In our implementation, we chose this threshold $\mu = 4$dB.

[3]In our implementation, we chose $(\tau_x, \tau_y, \tau_z) = (0.12, 0.14, 0.27)$

who did not have prior knowledge of the details of RF-AR's implementation, techniques, or the environment setup. In all of our experimental trials, the subjects were tasked to use the headset to find fully occluded RFID tagged items that were hidden in different locations. We evaluated multiple versions of RF-AR with varied degrees of user interface complexity against a baseline(XAR).

- Baseline(XAR): Similar to past work [2]. No feedback is provided to the user until the target RFID is located
- RF-AR (SNR): Provides color cues based on RF measurements as described in §V-A
- RF-AR (ConfSNR): Provides both SNR-based color cues and the confidence-based ellipsoid hologram as described in §V-A, and §V-B
- RF-AR (ArrowConf): Highest degree of complexity in user interface. Optimizes users' trajectory through visualizing arrows, as described in §V-C, as well as color cues and ellipsoid hologram

We randomized the order of baseline and RF-AR implementations when we asked each subject to use the headset.

Metrics: We evaluated RF-AR performance through four main metrics: *1) Localization error* is the error between the target RFID tag ground truth location and the the headset's estimated location of the target RFID. *2) Time* is the duration that the user spent searching from when they started the experiment until when the system found the target RFID tag confidently. *3) Success Rate* is the rate that headset was able to locate the target RFID tag within 300 seconds and with less than 25 cm of localization L2 norm error. *4) Traveled Distance* is the distance traveled by the user from the starting point until the RFID tagged item is confidently located.

Ground Truth: The investigators, who knew where the target RFID tags were hidden, dragged and aligned a holographic spheres in the Hololens App onto the target RFID tag location and sent the location to the edge server. The holographic spheres were removed before the headset was given to the subjects to ensure the RFID locations are unknown to subjects.

VIII. Results

We conducted 80 trials with 20 users to evaluate the impact of RF-AR's RF-based dynamic user interface and path optimization on the efficiency and accuracy of RFID localization.

A. Time Efficiency and Localization accuracy

We first analyze localization error as a function of the time it takes users to localize a target RFID tagged item. Fig.4a plots the median of L2 norm localization error against the median time that the users spent searching for the item over all trials. The red dot demonstrates the result for Baseline(XAR), pink dot shows RF-AR (SNR), and blue and green dots show RF-AR (ConfSNR) and RF-AR (ArrowConf), respectively. As shown in Fig.4a, Baseline(XAR) achieves a median localization error of 9.9 cm within a median time of 118 sec. RF-AR (SNR), RF-AR (ConfSNR) and RF-AR (ArrowConf) achieve median localization errors of 10.0 cm, 9.4 cm and 8.6 cm, and a median time of 36 sec, 58 sec, and 76 sec, respectively. We make the following remarks:

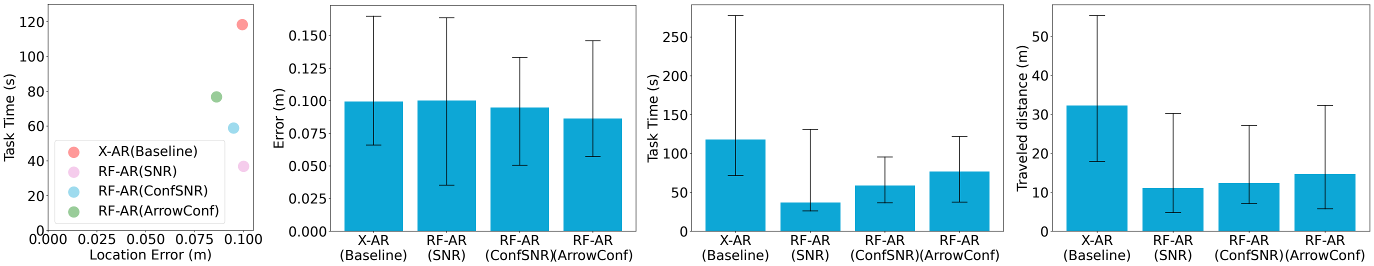
- All three versions of our system outperform the baseline in the median time. This proves that at a fundamental level, adding the"human-in-the-loop" is the key driver for performance improvements (reducing the time needed for localizing fully occluded RFID tagged targets) and delivers meaningful advantage over prior art.
- For the three versions of RF-AR, as the complexity of the user interface increased the median localization error decreased. These results demonstrate that increasing the UI cues and interface sophistication successfully influence the user's trajectory and resulted in higher quality RF measurement and improved localization accuracy.
- For the three versions of RF-AR, as the complexity of the user interface increased, the median time for localization increased. This demonstrates that while all variations of our design outperform the baseline in time, increasing the complexity of the user interface makes it *slightly* more time consuming for the users. This could be because more sophisticated prompts require longer reaction times from the general AR user.

For every user interface, we also independently analyzed the localization error, as shown in Fig.4b. The bars show the localization L2 norm error, and the error lines demonstrate the 10th and 90th percentile. In addition, we further analyze the time it takes users to finish the task in Fig. 4c, where bars show the median task times and error lines again indicate the 10th and 90th percentiles for each UI. Based on these two figures, we make the following remarks:

- Because Baseline(XAR) does not provide any feedback to the user, the 90th percentile of convergence time is 280 sec or approximately 4.6 minutes, which is over 2x the median time (118 sec). This clearly shows that not delivering any prompt to the user adversely affects the efficiency and reliability of the system.
- While RF-AR (SNR) has the lowest median time to finish the tasks (40 sec), its 90th percentile is 131 sec and is higher than the 90th percentile of the more sophisticated RF-AR (ConfSNR) and RF-AR (ArrowConf), which are 95 sec and 121 sec respectively. This demonstrates that RF-AR (SNR) is less reliable in improving the efficiency and accuracy of RFID tagged items' localization.
- More sophisticated user interfaces, RF-AR (ConfSNR) and RF-AR (ArrowConf), achieved 90th percentile localization error of 13.3cm and 14.6cm. While our simplest user interface, RF-AR (SNR), and the baseline have a 90th percentile localization error of 16.4cm. This demonstrates that although RF-AR (SNR) had a better median task time than more sophisticate version of RF-AR, it has less reliable localization accuracy in 90th percentile.

B. Success Rate

We define a trial to be a success when the localization error and time to find the item are below 25 cm and 300 seconds, respectively. We expect that an error above 25 cm



(a) Task Time vs Error (b) L2-norm Localization Error (c) Total Task Time (d) Total Traveled Distance

Fig. 4: **Evaluation Results.**(a) The plot shows the median L2-norm localization error and median task duration for X-AR (Baseline) in red, RF-AR (SNR) in pink, RF-AR (ConfSNR) in blue, and RF-AR (ArrowConf) in green. (b) The bar graph displays the median L2-norm localization error (c) The bar graph shows the median time that users took to complete the task (d) The bar graph shows the median distance that users traveled to complete the task. All error bars represent the 10^{th} and 90^{th} percentiles of the respective metrics.

would be insufficient granularity to distinguish between boxes on shelves. We also note that for trials lasting over 300 sec (5 min) users became discouraged to search. Table I reports the success rate for each UI over all trials. We note the following:

- Baseline(XAR) failed in 10% of the trials. All failures were due to users not finding the RFID tagged target item in time. This shows that the lack of user guidance can result in long search times that reduce the efficiency and user appeal of the system.

- RF-AR (SNR) also fails 10% of the trials. These failures were due to an instance of large localization error and an instance of long search time. This demonstrates that minimizing the cues to the user results in a higher likelihood of failure.

Methods	Baseline (XAR)	RF-AR (SNR)	RF-AR (ConfSNR)	RF-AR (ArrowConf)
Success Rate	90%	90%	100%	100%

TABLE I: **User Interface imapact on Success Rate**

C. Traveled Distance

We also evaluated the total distance the user has to travel (walk) from their starting position until they locate the hidden RFID tagged item. Fig.4d shows the users' traveled distance for each type of user interface. The bars show the median travel distance and the error lines demonstrate the 10th and 90th percentile. We make the following remarks:

- Baseline(XAR) needed a median of 32.22 m traveled distance to locate the RFID tagged item. This is significantly longer than other user interfaces. This proves that providing the user with real-time feedback based on the quality of RFID measurements greatly reduces the amount of walking or searching needed from the user.

- RF-AR (ArrowConf) required a median of 14.61 m traveled distance while RF-AR (SNR) and RF-AR (ConfSNR) needed 11.12 m and 12.40 m, respectively. The addition of the guiding arrow for DOP optimization in RF-AR (ArrowConf) results in a slight increase in the traveled distance. We believe that this is due to the arrow guiding the user to new locations to improve the confidence of the system and reduce the DoP to get lower the localization error.

IX. Conclusion

There have been a huge investments in AR/VR devices that show the potential to transform the way humans and technology interact. In this paper, we present a system that seamlessly integrates RF sensing with AR to enable users to find fully occluded RFID tagged items. We introduce a novel "human-in-the-loop" design that considers path optimization and presents a dynamic RF-based user interface. Our results show the potential of our human-centered design to improve the accuracy and efficiency of RFID localization for fast item retrieval, an application with important implications in sectors such as manufacturing, retail, and warehousing. In future implementations of RF-AR, the entire RF sensing hardware can be integrated into the AR headset.[4] As the research evolves, we envision future designs to explore new ways to synergistically combine wireless sensing and AR, and we hope that our work inspires further research in this space.

Acknowledgments We thank the anonymous reviewers, and the Signal Kinetics group for their help and feedback. This research is sponsored by NSF (Awards #1844280 and #2044711), the Sloan Research Fellowship, and MIT Media Lab.

References

[1] Tara Boroushaki, Laura Dodds, Nazish Naeem, and Fadel Adib. Fusebot: Rf-visual mechanical search. *Robotics: Science and Systems 2022*, 2022.

[2] Tara Boroushaki, Maisy Lam, Laura Dodds, Aline Eid, and Fadel Adib. Augmenting augmented reality with non-line-of-sight perception. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*, 2023.

[3] A. R. Chatzistefanou, G. Sergiadis, and A. G Dimitriou. Tag localization by handheld uhf rfid reader with optical and rfid landmarks. *IEEE JRFID*, 2023.

[4] Shing H. Doong. A closed-form formula for GPS GDOP computation. *GPS Solutions*, 13(3):183–190, July 2009.

[5] Sebastian Kunkel, R. Bieber, M. Huang, and M. Vossiek. A concept for infrastructure independent localization and augmented reality visualization of rfid tags. In *2009 IEEE MTT-S International Microwave Workshop on Wireless Sensing, Local Positioning, and RFID*.

[6] Zulqarnain Rashid, Enric Peig, and Rafael Pous. Bringing online shopping experience to offline retail through augmented reality and rfid. In *2015 5th International Conference on the Internet of Things (IOT)*.

[7] TeamViewer. https://www.teamviewer.com/en-us/frontline/xpick/, 2022.

[8] Lei Xie, J. Sun, Q. Cai, C. Wang, Jie Wu, and . Lu. Tell me what i see: Recognize rfid tagged objects in augmented reality systems. In *Proceedings of the 2016 ACM UbiComp*.

[9] Lei Xie, C. Wang, Y. Bu, J. Sun, Q. Cai, J. Wu, and S. Lu. Taggedar: An rfid-based approach for recognition of multiple tagged objects in augmented reality systems. *IEEE Transactions on Mobile Computing*.

[10] Akihiro Yamashita, K. Sato, S. Sato, and K. Matsubayashi. Pedestrian navigation system for visually impaired people using hololens and rfid. In *2017 Conference on TAAI*.

[11] J. Zhang, S.K. Ong, and A.Y.C. Nee. Rfid-assisted assembly guidance system in an augmented reality environment. *International Journal of Production Research*, 49(13):3919–3938, 2011.

[4]While the current prototype uses BladeRFs and a Raspberry Pi, future designs can leverage compact RFID reader chips like Lepton3 (∼1"x1"x0.1").