

Improving Augmented Reality Relocalization Using Beacons and Magnetic Field Maps

Niranjini Rajagopal, John Miller, Krishna Kumar Reghu Kumar, Anh Luong, Anthony Rowe

Electrical and Computer Engineering Department

Carnegie Mellon University

Pittsburgh, U.S.A.

{niranjir,jmiller4,kregghuku,anhluong2,agr}@andrew.cmu.edu

Abstract—In this paper, we show how beacon-based indoor localization and additional environmental fingerprints like magnetic field data can be used to both accelerate and increase the robustness of Augmented Reality (AR) relocalization. We show how the combination of Visual Inertial Odometry (VIO) and beacons can be used to construct a dense indoor magnetic field map that can act as a fine-grained calibration for compasses to quickly determine a mobile device's orientation. Unique to our approach is that we leverage accurate VIO trajectories to provide full vector magnetic field mapping that can be collected and used with devices placed in any orientation. We demonstrate a system running on an iPhone that can acquire location with 80th percentile 3D accuracy of 27cm in LOS and 46cm in NLOS, and our magnetic field mapping approach can instantly estimate orientation with 80th percentile accuracy of 11.7 degrees. We demonstrate an end-to-end system for generating a magnetic field map and subsequently localizing and tracking mobile users using beacons and VIO. This has the side effect of enabling multi-user (even cross-platform) AR applications, as all users can now be localized with respect to a common global reference without any sharing of visual feature maps.

I. INTRODUCTION

Recent advances in VIO and Simultaneous Localization and Mapping (SLAM) on devices ranging from headsets to smartphones have made AR an easily accessible commodity on an exciting number of platforms. Mobile AR APIs like those found in Apple's ARKit and Android's ARCore currently provide (1) VIO-based tracking, (2) scene understanding, and (3) realistic rendering based on features and lighting from the environment. This can support applications like gaming and product visualization, where the user interacts with the virtual objects for a single session. Advanced mobile AR and platforms such as Microsoft HoloLens and Magic Leap allow virtual content to *persist* across sessions (power cycles of the device) and across users (if the stored features are shared). Persistence of data on mobile devices opens up a variety of new applications where users can now, for instance, annotate items in large areas with virtual signage and navigation instructions, or provide contextual control panels for things like projectors and automation equipment.

Such persistence of virtual content is enabled by a relocalization process. We refer to the process of estimating the six-degrees-of-freedom pose estimation with respect to a

previously acquired map as relocalization. Relocalization is performed by comparing visual features in the environment obtained through the camera with previously stored features.

Unfortunately, an approach relying on vision has limitations when visual features are poor. As a result, in many cases, this relocalization process either takes an extended period of time or fails due to lack of texture, changes in lighting, or changes in the location of objects in the environment, like moving furniture or people. In cases where vision struggles, we advocate using localization infrastructure of range-based beacons which do not drift over time, are robust to environment dynamics, and are based in a global reference frame that is not device- or session-dependent.

While range-based beacons can aid in location acquisition, they do not provide sufficient information to acquire orientation. Orientation acquisition outdoors can easily be accomplished using a magnetometer; since the earth's magnetic field generally points in a constant direction outdoors, it can be used as a global orientation reference. However, indoor magnetic fields tend to fluctuate wildly across space, due to the metallic materials inside the building structures, as well as the objects inside. We show that with accurate 6DOF tracking from VIO, we now have enough information to map the 3D magnetic field vector at a fine enough granularity that it can calibrate orientation for future users. Existing works that have constructed such maps of the magnetic field direction typically employ a robotic system with the sensor mounted at a known orientation, and rely on accurate wheel odometry. These techniques are not practical for building scale, due to the cost and effort involved. Leveraging accurate VIO, our approach is the first that is able to determine the full vector magnetic field map such that it can be used to calibrate a mobile phone or headset compass held in any orientation. Our system maps the magnetic field by having users simply walk around wearing or holding the device normally. Subsequently, devices can instantly localize themselves using the beacons and instantly estimate their orientations using the on-board magnetic sensor and the previously obtained magnetic field map to drastically reduce vision-based search uncertainty.

We demonstrate the accurate location and orientation obtained by our system with AR applications that do not rely on visual point clouds for mapping. This is useful when the

environment is dynamic, crowded, or has low lighting, where visual mapping can fail, or in environments where an indoor localization system is already in place. Hence, there is no need for storing or sharing the visual point cloud for relocalization. In the future, our approach could easily be integrated with low-level visual relocalization to support areas with sparse or no beacon coverage. Given a high-confidence location and orientation, the search space for visual relocalization can be dramatically reduced and the tolerance for matches can be relaxed. Though demonstrated on a mobile AR platform, this same approach would easily apply to headsets or similar localization platforms that require full pose information.

The main contributions of this paper are:

- 1) A scheme where range-based beacons can provide fast and robust relocalization.
- 2) A method for rapid pose acquisition that uses crowd-sourced vectored (pose invariant) magnetic field maps. We experimentally show the performance of this approach across a variety of environments and over time.
- 3) An end-to-end system implementation of an accurate location and orientation acquisition system using beacons and magnetic field, demonstrated with multi-user AR application.

II. BACKGROUND AND RELATED WORK

In this section, we first discuss emerging range-based beacon technologies and motivate beacons for improving mobile AR. We then describe the state-of-art approaches for relocalization using vision. Finally, we discuss approaches that use magnetic field for location and orientation estimation.

A. Beacon ranging technologies

WiFi 802.11mc, which supports round-trip-time-of-flight ranging [1], is currently implemented in Android Pie and is rolling out in some of the newer WiFi Access Points. Ultrasonic beacons can provide accurate localization [2] on unmodified mobile devices. Since they operate at the speed of sound, their time-synchronization requirements can be met using RF signals such as 802.15.4 or BLE [3]. Ultra-Wideband (UWB) ranging is commonly used in localization system implementations due to readily available chipsets and is reducing in power with emerging standards such as 802.15.4z. Bluetooth Low Energy ToF ranging [4] holds promise as a highly prevalent ranging technology for peer-to-peer devices. We advocate leveraging range-based beacons as these emerging technologies are finding their way into commodity devices that are used and deployed indoors.

B. Vision-based relocalization

Relocalization, also referred to as pose acquisition or registration, using vision based approaches is performed by matching features in the field of view with pre-mapped visual features. Several research studies have shown accurate relocalization using machine learning techniques [5], [6] and by fusing additional sensing modalities such as WiFi and

magnetic field [7]. Most of these approaches have demonstrated accurate and efficient relocalization on datasets, but are computationally intensive for mobile devices. However, ARKit 2 by Apple and ARCore by Google have shown persistent AR by performing relocalization using pre-stored maps. Vision-based relocalization is promising, since it does not rely on external infrastructure. However, vision will always suffer in environments devoid of features, with changes in lighting, and when the scenery changes over time. In addition, vision-based relocalization in large areas requires searching through many candidate feature matches, which can become expensive if an initial location estimate is not provided. It is often the case that the user must walk around and view several areas of a scene before visual relocalization is able to take effect. We are optimistic that vision will continue to improve, but that there are certain environments like office cubicles, hospitals, or parts of airport terminals where even humans have trouble figuring out their location without exploring. Due to the limitations of a purely vision-based approach, we advocate combining visual approaches with range-based beacons.

C. Magnetic field sensing

Several prior works leverage the spatial variation and temporal stability of the magnetic field inside buildings for estimating location [8]–[12]. These approaches map the magnetic field magnitude as a function of location during the system setup and subsequently use it as a signature to estimate location. Since the magnetic field signature is not unique, it is integrated with other sensors or matched against a time series pattern as the user walks around [8], [10], [13], [14]. Our work differs from these approaches since we localize using beacons, and use the mapped magnetic field direction to estimate orientation.

Prior work has shown that orientation can be estimated using magnetic field by looking for opportunities when the change in orientation from the magnetometer is consistent with the gyroscope, indicating that the field is stable in that region [15]. The authors found on average two opportunities per minute when the field is stable. This requires the user to walk around and is not suitable for instant orientation acquisition. Robotic systems with sensors mounted in a known fixed orientation and accurate odometry use the magnitude and the direction of the magnetic field [16], but robots are neither affordable nor flexible enough for all applications. The closest to our work is [17], which presents a pedestrian-held system for magnetic field mapping. However, it relies on the phone being rigidly held and carried along a specific pre-determined route through the room. Our approach obtains the vectored magnetic field maps in any indoor space with a phone, where the user is free to walk through the room at any speed and any walking pattern.

III. SYSTEM OVERVIEW

The relocalization problem is equivalent to the problem of estimating the six-degrees-of-freedom pose of the device in an external fixed frame. Upon startup of an AR app, the AR

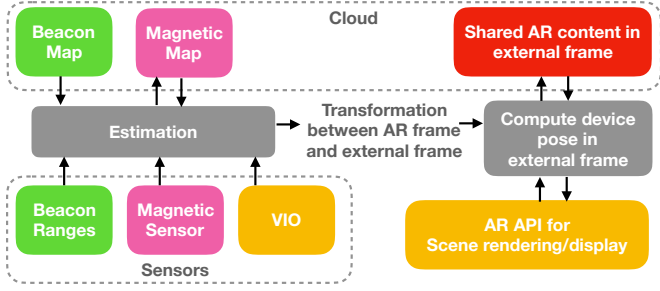


Fig. 1. System architecture for fusing beacon ranges and magnetic field sensor data to improve AR

session assumes a reference frame (AR frame) and the pose is locally tracked in this reference frame. To relocalize, we estimate the transformation between the AR frame and beacon frame (external fixed frame) using the following inputs: (1) range from beacons, (2) magnetometer, and (3) pose tracking using visual inertial odometry (VIO) from the AR API. In addition, we use a map of the beacon locations and a map of the magnetic field, which is updated continuously as the system is used. Our architecture is shown in Figure 1.

Range from beacons: Our approach uses time-of-flight distances or range measurements from beacons. We use the map of beacon locations as the external fixed reference frame, which we refer to as the beacon reference frame F_{BEAC} .

Magnetic Field: We use the magnetometer data on the mobile device for acquiring the orientation of the device. This sensor measures the 3D magnetic field in a frame of reference that is fixed with respect to the device, F_{DEV} .

Visual Inertial Odometry: Apple’s ARKit and Android’s ARCore APIs provide motion tracking by fusing tracked visual feature points with inertial sensor motion data, referred to as visual inertial odometry (by ARKit) or concurrent odometry and mapping (by ARCore). The output of VIO is the position and orientation of the device with respect to the AR reference frame at startup, F_{AR} . Next, we characterize various parameters of ARKit’s VIO tracking accuracy under different environments.

Drift: Figure 2(a) shows a 200m trace of VIO, where the overall drift is 5m. The drift is due to cumulative errors in the angle estimates, which are higher when the user takes turns while walking. To account for this drift, we assume the error in angle between two position updates separated by time δt is drawn from $\mathcal{N}(0, \delta t \cdot \sigma_\theta^2)$ where σ_θ is empirically evaluated.

Distance estimation: Figure 2(b) shows the VIO tracking over a rectangular trace of perimeter 100m in an office environment. We observe that the VIO traces are scaled to different extents as compared to the ground truth trace (shown by the black dots). This error in scale is due to incorrect distance estimation, which often occurs when a single camera is used and depth cannot be estimated correctly. Across all the experiments we conducted in real-world settings with trace lengths ranging from 10m to 300m, we estimated the error in VIO distance estimation, shown in Figure 2(c). This data implies that when the position changes by δd , the distance as estimated

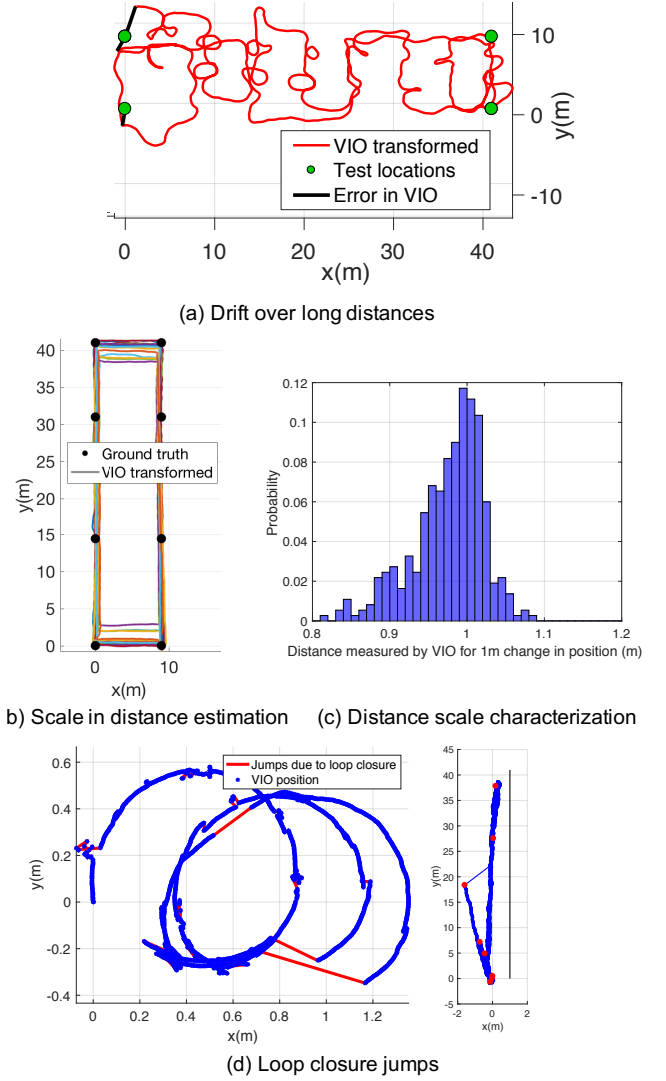


Fig. 2. Visual inertial odometry characterization

by VIO is drawn from a distribution $\mathcal{N}(0.98\delta d, (0.05\delta d)^2)$.

Errors in relocalization: Figure 2(d) shows two scenarios where there are abrupt jumps in position due to incorrect relocalization or loop closure. In the scenario on the left, we placed a phone on a cart and rotated the cart without any translation movement. In the scenario on the right, we walked through a hallway facing identical cubicles in periodic intervals and saw jumps due to different physical locations being visually identical. However, we do not observe abrupt jumps in longer traces, since there is a limited time-memory for the features and relocalization requires the features in memory to match currently visible features. This also highlights some of the challenges and limitations of heavily vision-based localization.

IV. ESTIMATION

In this section, we describe the *Estimation* block introduced in Figure 1. The estimation block has two outputs: (1) the transformation between the beacon and AR frame of the mobile device, which together with the VIO tracking gives the device’s pose in the external fixed frame, and (2) the magnetic

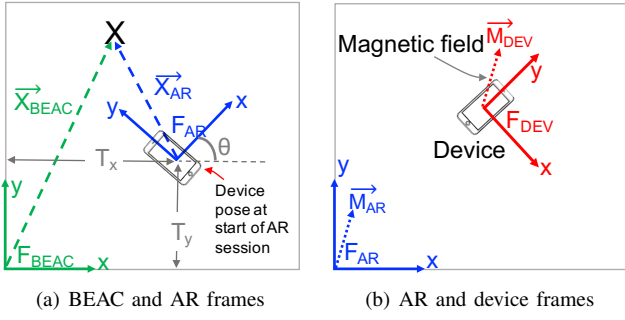


Fig. 3. Three frames of references (F_{BEAC} : Beacon frame; F_{AR} : AR frame; F_{DEV} : Mobile device frame)

field map. The sensor data used for estimation are collected in three different frames of reference: F_{BEAC} (ranges), F_{AR} (VIO), and F_{DEV} (magnetic field), shown in Figure 3. F_{BEAC} is the external fixed frame which is defined by the beacon locations. When a mobile device renders AR content, the AR Scene Rendering API renders the content in F_{AR} with coordinates X_{AR}^{obj} . For persisting or sharing the rendered content, we convert the content to the beacon frame F_{BEAC} to X_{BEAC}^{obj} and store its coordinates (or share with another user). When the device renders content retrieved from a previous session or another user, X_{BEAC}^{obj} , it converts it to F_{AR} as X_{AR}^{obj} before sending it to the AR API for rendering. Hence we have to estimate the six-degrees-of-freedom transformation between F_{AR} and F_{BEAC} . F_{AR} and F_{BEAC} have one axis parallel to the gravity vector. Though this is provided to us by ARKit, if that is not the case, we can compute the gravity vector by filtering the accelerometer data when the phone motion is low [15]. Having this axis aligned, we reduce the six-degrees-of-freedom frame conversion problem to a three-degrees-of-freedom translation, $T_{AR \rightarrow BEAC}$ and a one-degree-of-freedom rotation, $\theta_{AR \rightarrow BEAC}$. Figure 3(a) shows the translation in 3D and the rotation.

$$X_{BEAC} = R(\theta_{AR \rightarrow BEAC}) \cdot X_{AR} + T_{AR \rightarrow BEAC} \quad (1)$$

where

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

and

$$T_{AR \rightarrow BEAC} = [T_x, T_y, T_z]^T \quad (3)$$

Ideally, if there is no drift in VIO and no noise in VIO or range measurements, $(T, \theta)^{AR \rightarrow BEAC}$ would be time-invariant. However, in reality, the transformation is time-varying, and for accurate rendering we have to continuously update

$(T, \theta)^{AR \rightarrow BEAC}(t)$. In our state estimation, we are estimating $X_{BEAC}(t)$ and $\theta(t)$. We read X_{AR} from the VIO API and use these quantities to compute $(T, \theta)^{AR \rightarrow BEAC}$.

We chose a Particle Filter (PF) approach for our state estimation for the following reasons: (1) it is an online approach and we are only required to maintain the previous state at any

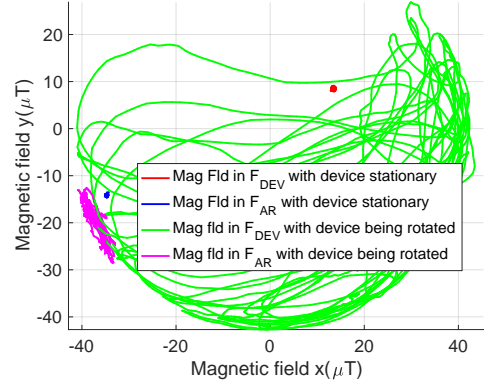


Fig. 4. Conversion of magnetic field from device frame F_{DEV} to AR frame F_{AR}

time (2) it allows us to accommodate for arbitrary noise models used to describe VIO and beacon errors, (3) it can continue to work in under-defined cases with as few as 1-2 beacons, and (4) it is agnostic to update rate and allows us to handle asynchronous ranges from the beacons without requiring us to receive synchronous ranges and perform trilateration.

The device state at time t is represented by N 5-dimensional particles (x, y, z, θ, s) , with weights w_t^i , $i \in [1, N]$, where (x, y, z) is the 3D position of the device in F_{BEAC} , $\theta_{AR \rightarrow BEAC}$ is the rotation of AR frame with respect to world frame, and s is the scaling factor of AR compared to true distance.

The general PF algorithm has four functions: initialization, motion model update, measurement update, and resampling [18]. The states are updated from VIO measurement (X_{AR}) and the weights are updated from the range measurement. During initialization, the location (x, y, z) is initialized from the range measurement. The angle $\theta_{AR \rightarrow BEAC}$ is initialized from the magnetic map value at location (x, y, z) . With motion inputs, each particle gets updated in the direction indicated by its belief of $\theta_{AR \rightarrow BEAC}$. Eventually, particles with $\theta_{AR \rightarrow BEAC}$ close to the true value will be closer to the true location and hence weighed higher when a range measurement is received. As per Equation 1, we estimate $T_{AR \rightarrow BEAC}$ from $\theta_{AR \rightarrow BEAC}$, X_{BEAC} which we get from the states x, y , and from X_{AR} , which is the location of the device in F_{AR} got by reading the VIO data.

We use the estimated device pose to also create and update the magnetic field map. The magnetic field map at all locations is initialized to a uniform distribution in $[0, 2\pi]$. Eventually, when the user moves, the particle filter converges and we use the estimated $\theta_{AR \rightarrow BEAC}$, or the rotation between the AR and beacon frames to build the map, and the estimated locations to build the map. Below, we describe the magnetic field mapping process after $\theta_{AR \rightarrow BEAC}$ has converged.

Step 1: Reading the magnetic field sensor: First, we read the 3D magnetic field M_{DEV} in the device frame of reference F_{DEV} . This frame is fixed with respect to the device as shown in Figure 3(b).

Step 2: Converting the field from device to AR frame:

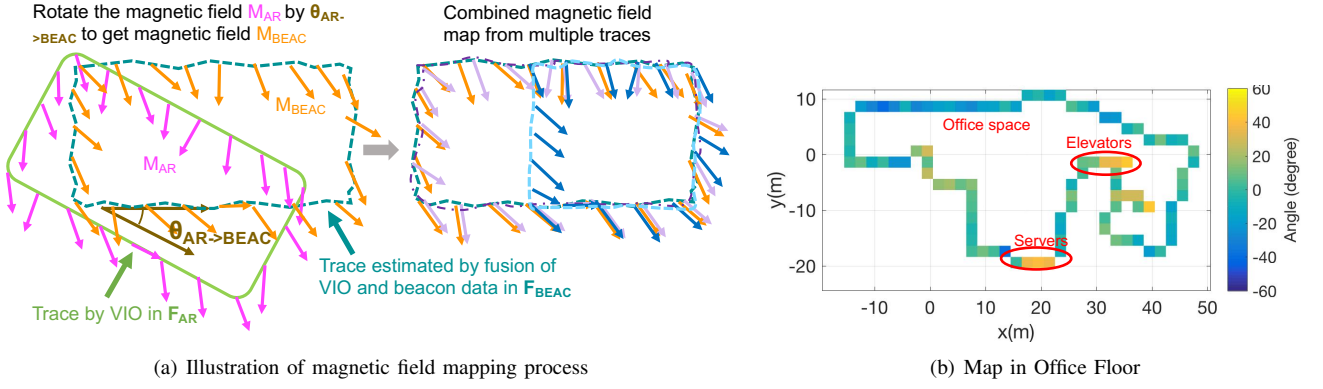


Fig. 5. Illustration of magnetic field mapping process

When the device moves, the AR frame of reference (F_{AR}) remains constant but the device frame of reference (F_{DEV}) changes. We transform the 3D magnetic field in device frame M_{DEV} to AR frame M_{AR} using the orientation, which is tracked by VIO. The result of this process is shown in Figure 4. We see that when the device is rotated about 360° in all orientations, the magnetic field measured in the device frame rotates about a sphere but the magnetic field in the AR frame is fixed.

Step 3: Converting the field from AR to beacon frame: We convert the magnetic field from the AR frame F_{AR} to the beacon frame F_{BEAC} using the estimated $\theta_{AR \rightarrow BEAC}$, which provides the rotation between the two frames of references about the z axis. We project the magnetic field M_{AR} to 2D on the $x-y$ plane and then rotate by $\theta_{AR \rightarrow BEAC}$ to compute M_{BEAC} . This process is illustrated in Figure 5(a).

Step 4: Building the magnetic field map: The magnetic field mapping process generates a map of the 3D magnetic field vector in the beacon frame. However, we store the map as a function of 2D locations to make the mapping process more tractable. We justify this choice in Section V-A1, by showing the impact of using a 2D map rather than a 3D map. As the system is in use, we obtain multiple traces from different users that span the same locations. The map at a location (x, y) is represented as the circular mean angle $M^\mu(x, y)$ and circular standard deviation angle $M^\sigma(x, y)$ of all the magnetic angles logged at location (x, y) . Figure 5(b) shows the magnetic field map of an office floor.

Orientation acquisition from the magnetic field map: Once the magnetic field is mapped, we use the map for orientation acquisition. When the particles are initialized, we initialize $\theta_{AR \rightarrow BEAC}$ from the map as:

$$\theta_{AR \rightarrow BEAC} \sim \mathcal{N}(M_{BEAC}(x, y) - M_{AR}, M^\sigma(x, y)) \quad (4)$$

In summary, upon startup, the device acquires its location from beacons, acquires its orientation from the magnetic field map, and as users walk around, the device pose is continuously tracked by fusion of beacon ranges and VIO, and the magnetic field map is continuously updated.

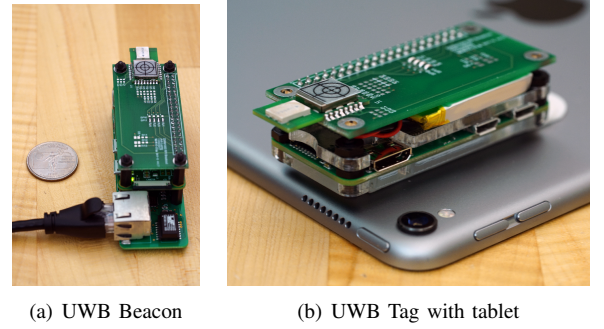


Fig. 6. Beacon and Tag Hardware

V. EVALUATION

To evaluate the system, we use Decawave UWB beacons for localization. However, any of the various emerging TOF ranging technologies could work. Decawave currently has a small UWB and BLE module [19] that can be easily attached to phones in order to range with UWB and provide position information over Bluetooth. In our prototype, we use a DWM1000 module connected to a Raspberry Pi Zero W shown in Figure 6. We evaluated our system in two environments. The first was a *Cafe* of size $145m^2$ ($1560ft^2$) which had 5 UWB beacons deployed in LOS of the test areas. The Cafe was located in a building with new construction and had large open areas with tables, chairs, counters, and pillars. The second was an *Office* area of size $400m^2$ ($4300ft^2$) with 10 beacons deployed in NLOS of the test areas. The office area had cubicles with office furniture and workstations. Test points and beacon locations were mapped accurately using a total station theodolite for ground truth. Four different users held devices in hand and walked at normal pace (4-5km/hr), clicking a button while crossing each test point.

A. Magnetic field feasibility

The magnetic field's spatial and temporal variation determines its feasibility for orientation acquisition.

1) *Spatial variation:* The spatial variation of the magnetic field along with the localization accuracy impacts the accuracy of the orientation acquisition. For instance, if the average localization error is $10cm$, and the magnetic field varies by

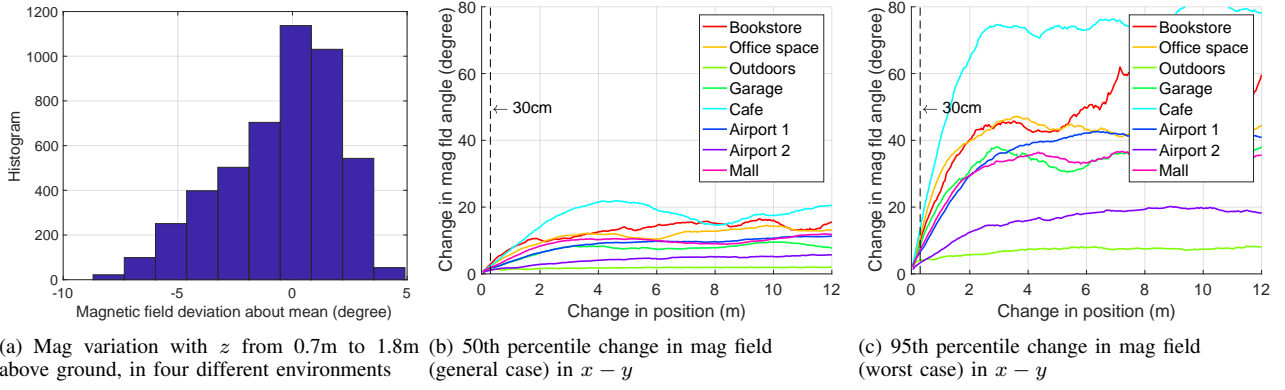


Fig. 7. Magnetic field spatial variation

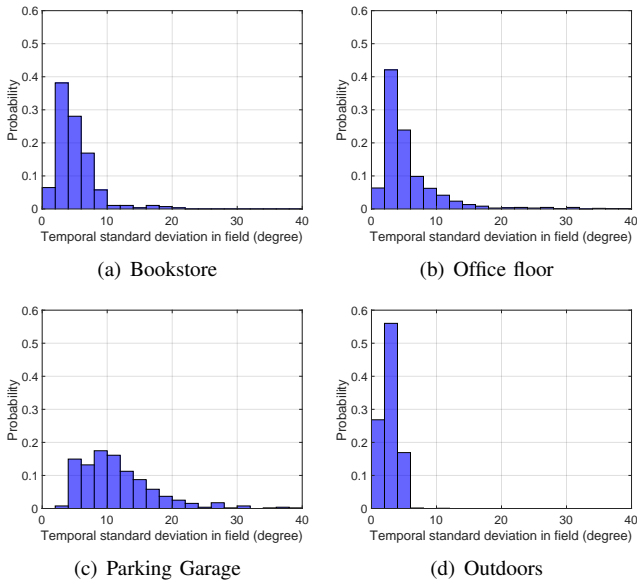


Fig. 8. Temporal variation of the magnetic field over 10 weeks. The plots show the histogram of the circular standard deviation across time for all test locations in the environment.

around 50° for locations that are 10cm apart, then this spatial variation of the magnetic field along with the localization error could result in a 50° error in orientation acquisition. To test the spatial variation, we experimentally evaluated the magnetic field variation along the vertical (parallel to gravity vector) and horizontal axes. Figure 7(a) shows the variation along the z -axis for device height of 0.7m to 1.8m above ground (1m about nominal standing height) across four environments (office, lobby, bookstore, cafe). Since the z -deviation is generally within 5° , we store the magnetic field map in 2D, rather than 3D, to keep the map size tractable without losing significant accuracy. Next, we characterize the variation over horizontal distance in several environments, including outdoors, garage, mall, cafe, and airports. We first compute the magnetic field angle difference between all pairs of locations that are a fixed distance apart. This captures the spatial variation for different distances. We then compute the CDF of the spatial variation for a fixed distance and extract the 50% spatial variation

(representative of generate case, shown in Figure 7(b)) and the 95% spatial variation (representative of worst case, shown in Figure 7(c)) for distances from $0 - 12\text{m}$. For instance, the *Cafe* line in Figure 7(c) shows that across all location pairs that are 1m apart, 95% have a field difference of less than 40° . As one might expect, with a larger change in distance, there is a higher change in field across all locations. However, beyond a certain distance, these changes are random. The change in magnetic field for a distance corresponding to 30cm (the approximate positioning error from our localization system shown in Figure 9(a)), is within 3° for 50% and varies from around $3 - 12^\circ$ for the different environments for 95%. This reinforces the notion that the magnetic field can be highly variable indoors and that certain measurements generalize well across recorded paths while other areas need dense recordings.

2) *Temporal variation*: Since orientation acquisition relies on matching with a previously acquired map, it is important to study the temporal variation of the field. Figure 8 shows the temporal variation over 10 weeks across four environments (book store, office floor, underground parking garage, outdoors). As expected, the outdoor environment has the least variation. The garage has the highest variation, due to cars (large metallic objects). In the two indoor environments, most regions have temporal standard deviations of less than 4° . In some environments, like stores with moving metal shelves, the magnetic field will often change after a reconfiguration. Through crowd-sourcing, our approach can continuously update and average magnetic field values from multiple users to adapt over time.

B. Localization accuracy

Figure 9(a) shows the localization accuracy of the phone in both environments, Cafe with LOS+NLOS, and Office with only NLOS, by fusion of beacon ranges and VIO. Users walked continuously between test points, stopping for at least one second at each test point and pressing a button on the app. Since the users held the device in hand above the test point marked on ground, this introduced $10 - 15\text{cm}$ error between the device's true location and the test point. In the Office environment and Cafe, we measure 80% error of 46cm and 27cm respectively.

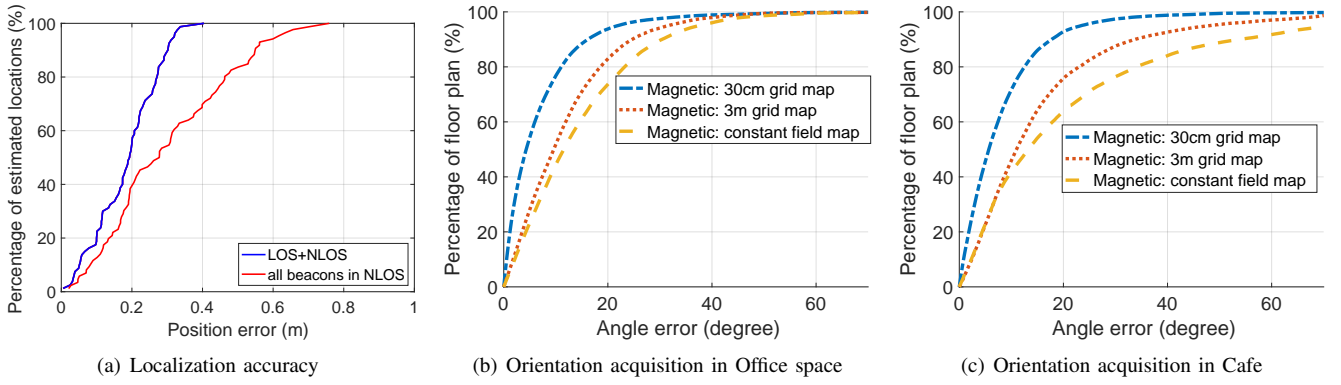


Fig. 9. Location and orientation acquisition accuracy

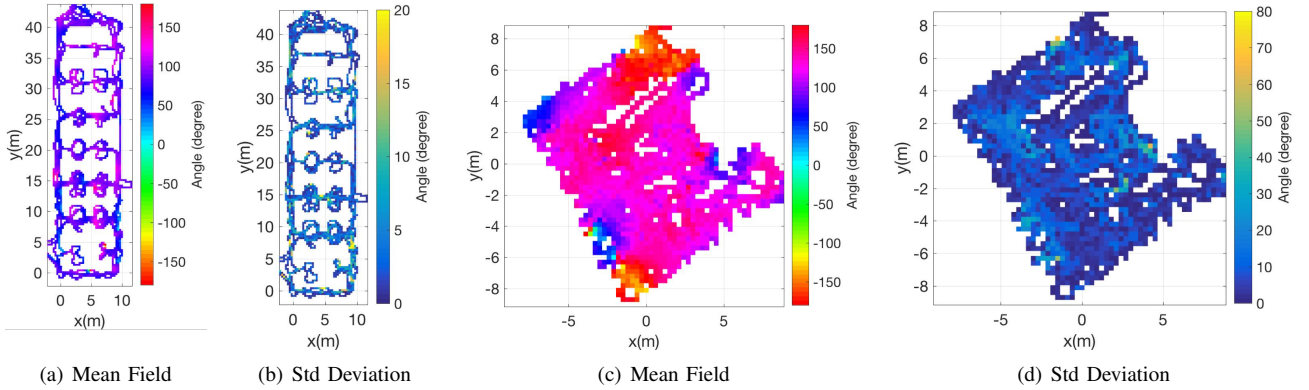


Fig. 10. Magnetic field map captured in the two deployments. Office environment (a-b) and a Cafe (c-d)

C. Magnetic field orientation accuracy

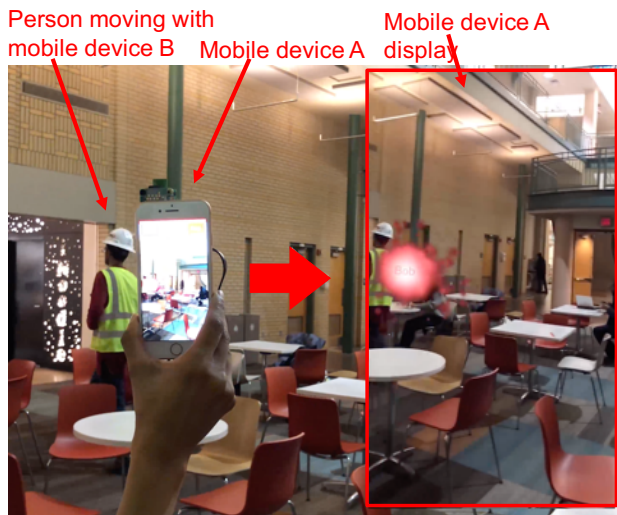
Figure 10 shows the magnetic field map in the two environments. The map is represented by the mean and standard deviation of the magnetic field. For the Office space, the standard deviation is 2° and 10° for 50% and 95% of the locations, respectively. For the Cafe, the standard deviation is 5.3° and 20.3° for 50% and 95% of the locations respectively for a grid size of 30cm. As shown in subsubsection V-A1, this Cafe has very high spatial variation compared to typical environments.

Figure 9 shows the accuracy of the magnetic-field based orientation for maps of varying spatial density. The *Magnetic: 30cm grid map* and *Magnetic: 3m grid map* show the angle error for two different grid sizes. The *Magnetic: constant field map* shows the performance if we assume the entire map has the same field, for which we used the mean of the field across the map. This is effectively equivalent to representing the map by a single sample. In the Office environment, we observe that the 80% point for 30cm is 12.5° and with a constant field is 30° . In the Cafe environment, we observe that the 80% point for 30cm is 11.0° and with a constant field is 54° . Across both environments, the 80% error is 11.8° for a 30cm grid size. We also see that the 3m grid performance is closer to assuming a constant field than using a 30cm grid which highlights the utility of our approach in areas with high spatial variation. To reduce map storage, regions of the environment

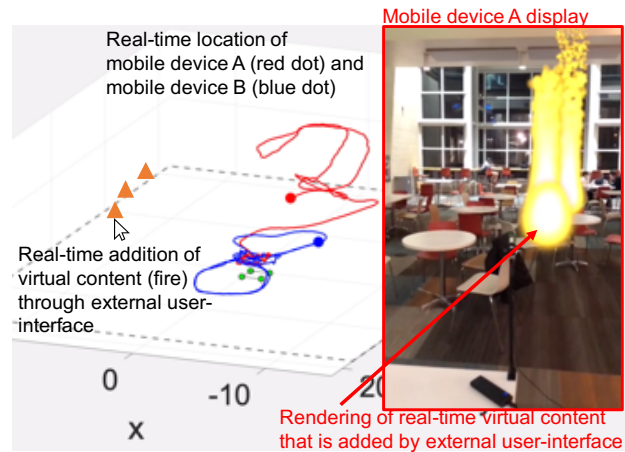
with high spatial variation should have a fine grid, while other regions can have a coarser grid. We also do not require complete coverage of a map, since a recorded trajectory can be extrapolated to support other nearby starting points with a confidence derived from the typical spatial variance captured in that area.

D. Demonstration application

We built the end-to-end system using UWB beacons and demonstrate it with an AR app, shown in Figure 11(a). Our setup has four beacons deployed at fixed locations. Two persons walk around with mobile devices attached to UWB tags. The devices are localized with respect to the beacons, and the real-time location pose information is available on an external console with a user interface. Figure 11(a) shows a find-a-friend application where both mobile devices are able to share persistent AR content. We see that mobile device A is able to see the mobile device B in real-time in AR (with a red blob located around the mobile device B). Figure 11(b) shows the location of both devices being available on an external console. We see that a user is able to click at locations on the user interface (shown as orange triangles) and the mobile device is able to see new virtual content (fire) pop up in real-time at the user-defined locations. This demonstrates the sharing of persistent AR content among the two mobile devices and the external system.



(a) Find-a-friend application



(b) Real-time location-based commands in AR

Fig. 11. AR demonstration application

VI. CONCLUSIONS AND FUTURE WORK

This work presented a proof-of-concept for using beacons integrated with VIO to provide persistent AR on mobile devices. One of our key contributions is that we can leverage VIO to capture the full pose of a mobile device allowing us to record the full magnetic field vector and not just magnitude. Though our system was designed for persistent AR, we can also apply this approach to creating other dense sensor maps for applications like indoor WiFi/cellular coverage and 3D scanning of spaces. In the future, we will explore ways to seamlessly integrate the beacon map and visual point-cloud maps so that the two systems can be combined for more effective relocalization.

VII. ACKNOWLEDGEMENTS

This work was supported in part by the CONIX Research Center, one of six centers in JUMP, a Semiconductor Research Corporation (SRC) program sponsored by DARPA, and by the NIST Public Safety Communications Research division.

REFERENCES

- [1] "https://www.wi-fi.org/news-events/newsroom/wi-fi-certified-location-brings-wi-fi-indoor-positioning-capabilities (viewed 10 Oct 2018)."
- [2] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Alps: A bluetooth and ultrasound platform for mapping and localization," in *The 13th ACM Conference on Embedded Networked Sensor Systems (SenSys 2015)*, ser. SenSys '15, 2015.
- [3] N. Rajagopal, P. Lazik, N. Pereira, S. Chayapathy, B. Sinopoli, and A. Rowe, "Enhancing indoor smartphone location acquisition using floor plans," in *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks*. IEEE Press, 2018, pp. 278–289.
- [4] "http://www.ti.com/lit/wp/sway008/sway008.pdf."
- [5] A. Kendall, M. Grimes, and R. Cipolla, "Posenet: A convolutional network for real-time 6-dof camera relocalization," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 2938–2946.
- [6] R. Clark, S. Wang, A. Markham, N. Trigoni, and H. Wen, "Vidloc: A deep spatio-temporal model for 6-dof video-clip relocalization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 3, 2017.
- [7] R. Clark, S. Wang, H. Wen, N. Trigoni, and A. Markham, "Increasing the efficiency of 6-dof visual localization using multi-modal sensory data," in *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*. IEEE, 2016, pp. 973–980.
- [8] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magical: Indoor localization using pervasive magnetic field and opportunistic wifi sensing," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1443–1457, 2015.
- [9] K. P. Subbu, B. Gozick, and R. Dantu, "Locateme: Magnetic-fields-based indoor localization using smartphones," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 4, no. 4, p. 73, 2013.
- [10] S. Wang, H. Wen, R. Clark, and N. Trigoni, "Keyframe based large-scale indoor localisation using geomagnetic field and motion pattern," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 1910–1917.
- [11] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. R. Choudhury, "No need to war-drive: Unsupervised indoor localization," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 197–210.
- [12] D. Caruso, A. Eudes, M. Sanfourche, D. Vissière, and G. Le Besnerais, "A robust indoor/outdoor navigation filter fusing data from vision and magneto-inertial measurement unit," *Sensors*, vol. 17, no. 12, p. 2795, 2017.
- [13] P. Robertson, M. Frassl, M. Angermann, M. Doniec, B. J. Julian, M. G. Puyol, M. Khider, M. Lichtenstern, and L. Bruno, "Simultaneous localization and mapping for pedestrians using distortions of the local magnetic field intensity in large indoor environments," in *International conference on indoor positioning and indoor navigation*. IEEE, 2013, pp. 1–10.
- [14] Q. Wang, H. Luo, F. Zhao, and W. Shao, "An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks," in *Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on*. IEEE, 2016, pp. 1–8.
- [15] P. Zhou, M. Li, and G. Shen, "Use it free: Instantly knowing your phone attitude," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 605–616.
- [16] H.-S. Kim, W. Seo, and K.-R. Baek, "Indoor positioning system using magnetic field map navigation and an encoder system," *Sensors*, vol. 17, no. 3, p. 651, 2017.
- [17] E. Le Grand and S. Thrun, "3-axis magnetic field mapping and fusion for indoor localization," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*. IEEE, 2012, pp. 358–364.
- [18] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005.
- [19] "https://www.decawave.com/products/dwm1001-module (viewed 1 Aug 2018)."