The Impact of Height on Indoor Positioning With Magnetic Fields

David Hanley, *Member, IEEE*, Augusto S. Dantas de Oliveira, Xiangyuan Zhang, *Member, IEEE*, Dae Hyun Kim, Yusheng Wei, and Timothy Bretl, *Senior Member, IEEE*

***Abstract*— Steel studs, heating, ventilation, and air conditioning (HVAC) systems, rebar, and many other building components produce spatially varying magnetic fields. Magnetometers can measure these fields and can be used in combination with inertial sensors for indoor positioning of robots and handheld devices, such as smartphones. Current methods of localization and mapping with magnetometers are often based on the simplifying assumption that magnetic fields do not vary with height. In this article, through the analysis of a large data set collected across three buildings on the University of Illinois campus, we quantify the extent to which this "planar assumption" is likely to be violated and examine the consequences for indoor positioning. First, we show that out-of-plane variations in the magnetic field were significant at over half the locations where magnetometer measurements were taken. Second, we show that absolute trajectory error in positioning was low when both localization and mapping were based on magnetometer measurements taken at the same height, but that error increased significantly with even small differences between these heights. Third, we show that the choice of height at which to take measurements—if this height was kept the same for both localization and mapping—had no significant impact on absolute trajectory error when averaged across a given set of trajectories although some trajectories existed for which different measurement heights led to significantly different errors. Fourth, we show that absolute trajectory error decreased when magnetometer measurements were aggregated across a small range of heights to produce a single, planar map and when measurements at the median height were used for localization.**

***Index Terms*— Gaussian process, indoor localization, magnetic field mapping, magnetic localization, particle filter.**

I. Introduction

SPATIALLY varying magnetic fields are produced by common building components and can be measured by magnetometers for indoor positioning. Magnetometers are small and inexpensive, are already carried by most robots and handheld devices, require no infrastructure other than the buildings themselves, and—when used with inertial sensors or wheel encoder odometry—result in positioning error that can be competitive with localization technologies that use Wi-Fi or Bluetooth [1]–[3]. These advantages have been seen as compelling enough in certain applications that some companies (e.g., IndoorAtlas and Sysnav) are now marketing indoor positioning solutions that are based on the measurement of magnetic fields.

Methods of indoor positioning that are based on the measurement of magnetic fields typically assume that these magnetic fields do not vary with height [2], [4]–[27]. This assumption is known to be reasonable when the magnetometer used to measure the magnetic field is approximately halfway between floor and ceiling in a building made only of vertically oriented ferromagnetic beams [28]. However, buildings are made up of more than vertically oriented beams, and magnetometer height can be well outside this range, depending on what or who is carrying the magnetometer and on how it is being held. The extent of the resulting variation in magnetic field strength and the impact of this variation on positioning accuracy cannot be derived from existing data sets, which have been generated either by taking measurements at a constant height or by aggregating all measurements, regardless of the height [29]–[32].

Our goal in this article is to critically examine the "planar assumption" and to quantify the impact of height on these existing methods of indoor positioning with magnetic fields. We restrict our attention to methods that are based on the use of planar maps, even if magnetometer measurements for mapping and localization are collected at different heights. In particular, we present the results of large-scale experiments in multiple buildings that help to answer four questions.

1) Where in a building does the magnetic field vary significantly with height? Our first task was to decide if the planar assumption is at all reasonable. To answer this question, we established a threshold for "significant variation" and computed the fraction of each building at which the difference between magnetometer measurements at different heights exceeded this threshold. We found that out-of-plane variations were significant at over half the planar locations where measurements were taken (see Section IV).
2) Does it matter if the magnetic field is measured at one height for mapping and another height for localization? One way in which significant out-of-plane variation in

Manuscript received November 16, 2020; revised January 5, 2021; accepted January 22, 2021. Date of publication February 12, 2021; date of current version March 4, 2021. This work was supported by the National Science Foundation under Grant 14-46765, Grant 14-27111, and Grant 15-44999. The Associate Editor coordinating the review process was Dr. Jesús Ureña. *(Corresponding author: David Hanley.)*

David Hanley, Xiangyuan Zhang, and Yusheng Wei are with the Department of Electrical and Computer Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801 USA (e-mail: hanley6@illinois.edu; xz7@illinois.edu; ywei26@illinois.edu).

Augusto S. Dantas de Oliveira, Dae Hyun Kim, and Timothy Bretl are with the Department of Aerospace Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801 USA (e-mail: dantasd2@illinois.edu; daehyun2@illinois.edu; tbretl@illinois.edu).

Digital Object Identifier 10.1109/TIM.2021.3059317

1557-9662 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

the magnetic field might have an impact on positioning accuracy is if measurements at different heights (e.g., from magnetometers inside differently sized robots) are used for mapping and localization.

To determine the extent of this impact, we computed the absolute trajectory error (ATE)—averaged across a set of reference trajectories in each building—as a function of this difference in height. We found that the error was low when both mapping and localization were based on measurements at the same height, but that error increased significantly with even small differences between these heights (see Section V).

3) Does it matter at which height the magnetic field is measured, if this height is kept the same for both mapping and localization? Another way in which significant out-of-plane variation in the magnetic field might have an impact on positioning accuracy—even if we keep measurement height the same for both mapping and localization—is if *in-plane* variation is more informative at some heights than at others (e.g., if putting a magnetometer at the top of a mobile robot works better than putting it at the bottom). To answer this question, we computed the ATE—again, averaged across a set of reference trajectories in each building—as a function of measurement height. We found that the choice of height had no significant impact on average error although some trajectories existed for which different measurement heights led to significantly different errors (see Section VI).
4) Does it matter if measurements of the magnetic field at different heights are aggregated to produce a single, planar map? So far, we have considered only planar maps that are constructed from measurements that are collected at the same height. In practice, it is also common to construct a single, planar map from measurements that are collected at many different heights (e.g., when surveying a building with a multirobot team, or when "crowd-sourcing" from smartphones carried by many different people). Our last question, therefore, is whether this process of aggregation leads to a change in positioning error that is similar to what we observed when mapping and localizing at different heights. As before, we averaged the ATE across a set of reference trajectories, in this case, expressing it as a function of the range of heights over which we aggregated measurements to produce a map. Surprisingly, we found that error decreased when the range of heights was nonzero but small, a result that we explain by analogy to map roughening (see Section VII).

Consistent with the experimental methodology described by the international standard for testing and evaluating positioning technologies [33], we collected the data on which these results were based by measuring the magnetic field at nearly two thousand locations within three different buildings on the campus of the University of Illinois at Urbana–Champaign. In particular, at each location, measurements were collected with 25 magnetometers that were spaced evenly from ankle height to head height and were carried by a wheeled mobile robot. The vertical separation between one magnetometer and the next was 7.62 cm, which is approximately one standard deviation of the height of men and women [34]. Existing methods of mapping [with Gaussian process regression (GPR)] and localization (with a particle filter to fuse data from inertial sensors) were used to draw conclusions about ATE (which we define in detail in the Appendix). These conclusions are of course only valid for the buildings and trajectories that we considered—our hope is that we, nonetheless, provide a foundation for more general conclusions to be established by other researchers in future work. As a supplement to this article, we have made all of our data and code freely available so that our results are both transparent and reproducible.[1]

Section II reviews prior work on localization and mapping with magnetic fields. Section III describes our experimental methodology. Sections IV–VII present our answers to each of the four questions posed above, in the order they are listed. Section VIII reflects on the implications of our results and highlights opportunities for future work.

A. Extensions to Our Own Prior Work

This article makes five extensions to our own prior work, which examined the magnetic field in smaller parts of two buildings and identified locations at which variations of the magnetic field with respect to height were significant [35]. First, we present an entirely new and publicly available data set, which we use for the analysis conducted in Sections IV–VII. Second, we present the new experimental apparatus that we built to collect this data set (see Section III). Some of the changes in this new apparatus include a different choice of magnetometers (with better resolution, lower noise, and lower hysteresis than those sensors used in our prior work), a larger number of these magnetometers, and a mechanical design that reduces the error in estimated height difference between magnetometers to less than 1 cm. Third, we present a new experimental methodology that we applied to collect our data set. Some of the changes in this methodology include the use of a visual-inertial navigation system that enabled the collection of data over larger areas, the calibration of misalignment between sensors, and the practice of stopping the robot at a sequence of points to reduce confounding factors. Fourth, while the question that we pose and answer in Section IV is similar to one we posed in our prior work, the method of analysis that we apply here is entirely new. Fifth, we pose and answer three new questions in Sections V, VI, and VII, which were not addressed in our prior work.

II. Related Work

Measurements of the geomagnetic field, for example, from a magnetic compass, have long been used to obtain estimates of heading for outdoor navigation [36]. In this context, the so-called "magnetic anomalies" that are associated with indoor environments—caused by the presence of steel studs, heating, ventilation, and air conditioning (HVAC) systems, rebar, and many other building components—are sources of

[1]Available on IEEE DataPort, DOI: https://dx.doi.org/10.21227/t881-758110.21227/t881-7581.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

disturbance that could reduce the accuracy of heading estimates [37]. Some methods have been proposed to identify and reject these sources of disturbance [38]. However, it has been more common simply to ignore magnetometer data when indoors [39].

Within the past 15 years, indoor magnetic anomalies have come to be viewed not as sources of disturbance but rather as features for mapping and localization. Most work in this area has focused on the problem of identifying the position of a robot or a person in two dimensions—so, with respect to the floor plan of a building—it has taken one of two general approaches to solve this problem [1]. The first approach is to create a planar map of the magnetic field as a function of floor position and to use a particle filter for localization [9]. Among the many variants of this first approach include those that incorporate other sensor data, such as wheel encoder odometry [2], that change the way in which particle weights are derived to reduce computation time [12], that take steps to enable simultaneous localization and mapping (SLAM)—rather than just localization—by identifying loop closure [40], and that use a Rao–Blackwellized particle filter to enable simultaneous localization and calibration [41]. This general approach to localization is also very similar to approaches taken by researchers localizing robots and pedestrians using measurements of Wi-Fi strength [42], [43]. The second approach is to create a set of fingerprints against which magnetometer measurements are matched [28]. Variants of this second approach have considered diverse features to use for the fingerprints [23], have explored the use of recurrent neural networks to learn these features from data [25], and have used convolutional neural networks to match fingerprints to magnetic maps [44]. Common to nearly all variants of both general approaches has been a choice to ignore out-of-plane variation in the magnetic field, something we refer to as the "planar assumption" [35].

Only recently has the community begun to explore the use of magnetic anomalies as features for indoor positioning in three dimensions. Most work in this area has focused exclusively on mapping, assuming that localization would proceed similar to what was described by Haverinen and Kemppainen [9] in two dimensions. In particular, Akai and Ozaki [45] proposed the use of GPR to create 3-D magnetic maps from magnetometer measurements. Recognizing both the merits of this approach and its significant computational demands—scaling quickly with the number of measurements that seem to be required in order to characterize a 3-D magnetic field—other researchers, such as Solin *et al.* [3] and Jidling *et al.* [46], have suggested variants that reduce computation time. Some effort has also been made to automate the process of collecting measurements in three dimensions throughout a building, for example, with an aerial vehicle [47]. Limited success has been reported so far on the use of 3-D magnetic maps for localization, with results coming mainly from alternative approaches—for example, a method of SLAM [48] or a method of visual-inertial odometry that incorporates magnetic measurements [49], [50].

In this article, we focus on the use of magnetic anomalies for 2-D indoor positioning. We do not propose a new positioning method—instead, we adopt the widely used method of Haverinen and Kemppainen [9], which suggests creating a map of the magnetic field based on magnetometer measurements collected offline and then applying a particle filter for localization based on magnetometer measurements collected online. As suggested by Akai and Ozaki [45], we assume that GPR is used to create the map. We restrict our attention to planar maps, given the lack of strong evidence so far that 3-D maps would reduce positioning error in two dimensions [1]. Our goal is to answer a number of questions (see Section I) about how these planar maps and the results of using them for localization are impacted by out-of-plane variation in the magnetic field, a topic that—as we have said—has been largely ignored in prior work [35].

III. Experimental Methodology

In this section, we say how we measured the magnetic field in three buildings. The resulting data are used in all subsequent sections of this article. Throughout this section and subsequent ones, confounding factors are identified and discussed as part of general descriptions of methodology. A summary of these confounding factors is provided in the Appendix.

A. Sensor Choice

We collected magnetic field data with three-axis RM3100 magnetometers from PNI Corporation. The PNI RM3100 magnetometers are magnetoinductive sensors, which incorporate a solenoid around a highly permeable magnetic core. These sensors operate on the principle that the inductance of the core will vary with the applied magnetic field [51]–[53].

The PNI RM3100 is both small (evaluation boards are $25.4 \times 25.4 \times 8.0$ mm) and cheap (evaluation boards are \$20 as of December, 2020). Therefore, it would not be an unusual choice for a magnetometer on ground robots or on mobile devices. Nonetheless, we acknowledge that most mobile devices, currently, use magnetoresistive magnetometers instead of magnetoinductive magnetometers, such as the PNI RM3100. Indeed, we also used a magnetoresistive magnetometer—the NXP MAG 3110—in our own prior work [35].

We chose to switch from magnetoresistive magnetometers in prior work to magnetoinductive magnetometers in the current work in order to improve the accuracy with which we can measure the magnetic field of a building. In particular, the PNI RM3100 experiences less hysteresis than the NXP MAG 3100 and is less sensitive to changes in temperature. The NXP MAG 3110 has a hysteresis of 0.25%, while the PNI RM3100 has a hysteresis of only 0.00375%. The NXP MAG 3110 has a temperature sensitivity of 0.1(%/°C), while the PNI RM3100 temperature sensitivity is not reported due to the fact that changes in core bias resistance or inductance due to temperature are experienced in both the forward and reverse directions of sensor oscillations and are, thereby, canceled.

We emphasize that different magnetometers may produce different results—different answers to the four questions posted in Section I—then what we report in the rest of this article. Our goal in using more accurate magnetometers (the

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

PNI RM3100 instead of the NXP MAG 3110) was not to make our results "better." Instead, our goal was to allow other researchers in future work to explore how results vary with sensor accuracy, a topic that is outside the scope of this article. In particular, it would be straightforward to simulate the use of less accurate magnetometers by artificially adding noise and reducing the sensitivity of measurements in our publicly available data set (but, of course, would be impossible to simulate the use of *more* accurate magnetometers than were used to construct our data set).

While we chose the PNI RM3100 magnetometer to limit the effect of temperature and hysteresis, this magnetometer still experiences a sensor random walk. The implication of this random walk is that, over time, the hard iron bias of all of our sensors will change. Since we are looking for differences in magnetic measurements, this effect serves as a potential confounding factor in our experiments that we cannot eliminate. Additional sensor details are described in the Appendix.

B. Ground-Truth Position and Orientation

We used Google Tango to produce a visual feature map and subsequently localize a Lenovo Phab 2 Pro attached to a ground robot. This device provided us knowledge of the position and orientation of each magnetometer in a scene. When a building lacks sufficient visual features, however, a system, such as the Google Tango, can fail to localize. There is existing literature that has identified and investigated solutions to this problem. For example, Deng *et al.* [54] plan the trajectory of a ground robot in order to keep a sufficient number of features in view of the camera at all times. DeGol *et al.* [55] adds AprilTags to several buildings to improve structure from motion average reprojection error, the number of registered points, and the number of registered images. Therefore, we added visual features (in the form of AprilTags) to the test area prior to mapping [56] so that there are always a sufficient number of features in view of our robot throughout our tests. We explored each area and provided multiple opportunities for drift in the map to be corrected through the identification of loop closures. Each map of the environment was saved in an area description file. That file was then used to localize our test devices. As a way to assess our ability to use Google Tango as a source of position and orientation measurements, we included a motion capture arena in a portion of one of our case study buildings: the Coordinated Science Laboratory (CSL). We then computed the ATE of the test points collected in the relevant portions of our buildings [57].[2] A histogram of the ATE for this case is shown in Fig. 1. The root mean square ATE in CSL is 5.9 cm.

C. Experimental Rig

Attached to a TerraSentia robot, we created a rig with 25, three-axis magnetometers spaced 7.62 cm apart vertically. The TerraSentia robot was initially developed and presented by Kayacan *et al.* [58]. The magnetometers were connected to five Arduino Megas, which were, in turn, connected to the robot's main computer. Motion capture markers and the Lenovo Phab 2 Pro were also attached to this rig. Fig. 2 shows that the robot and the attached rig on which magnetometers were mounted annotated with distance measurements.

[2]See the Appendix for a summary of how ATE is commonly defined.

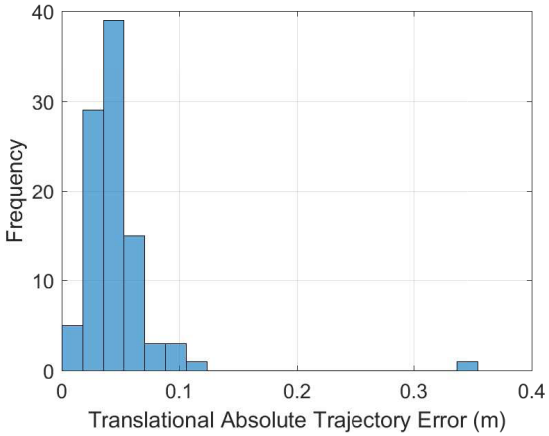


Fig. 1. Histogram of the ATE of the Google Tango in a portion of the CSL. ATE is listed in meters and is computed where motion capture data are available in our buildings.

D. Sensor Calibration

The magnetometers were calibrated together for hard and soft iron biases while attached to the robot. This calibration was done outside and away from buildings. Hard iron and soft iron biases were estimated using the procedure described by Merayo *et al.* [59]. In a constant magnetic field, plotting the measurements from a rotating, calibrated, triaxial magnetometer should produce a sphere whose radius is the magnitude of the constant magnetic field. This calibration found the hard and soft iron biases of our magnetometers by projecting data used for calibration to the unit sphere. To complete the calibration, we scaled our result by the magnitude of the magnetic field at Urbana, IL, USA, as given by the International Geomagnetic Reference Field (IGRF) model: 52.7076 μT. There are many alternative models of the magnetic field that could be used in our experiments. For example, the World Magnetic Model (WMM), the Enhanced Magnetic Model (EMM), and the High Definition Geomagnetic Model (HDGM). The HDGM requires a license to be purchased from the National Oceanic and Atmospheric Administration. The IGRF, WMM, and EMM are publicly available. In Urbana, the difference between these three models was less than 0.09 μT.

There are two types of misalignment that could have affected our experimental data collection. The first type of error (which we call intrasensor misalignment) is the error in measurement that results from a given RM3100 sensor triad not aligned in a mutually orthogonal manner. Note that if we calibrate for hard and soft iron bias, we also calibrate for these intrasensor misalignment errors as well [60]. The second type, intersensor misalignment, is the error in measurement resulting from differences in the axes between two different RM3100 sensor triads. After calibrating our sensors for hard and soft iron biases, we took multiple measurements of each

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

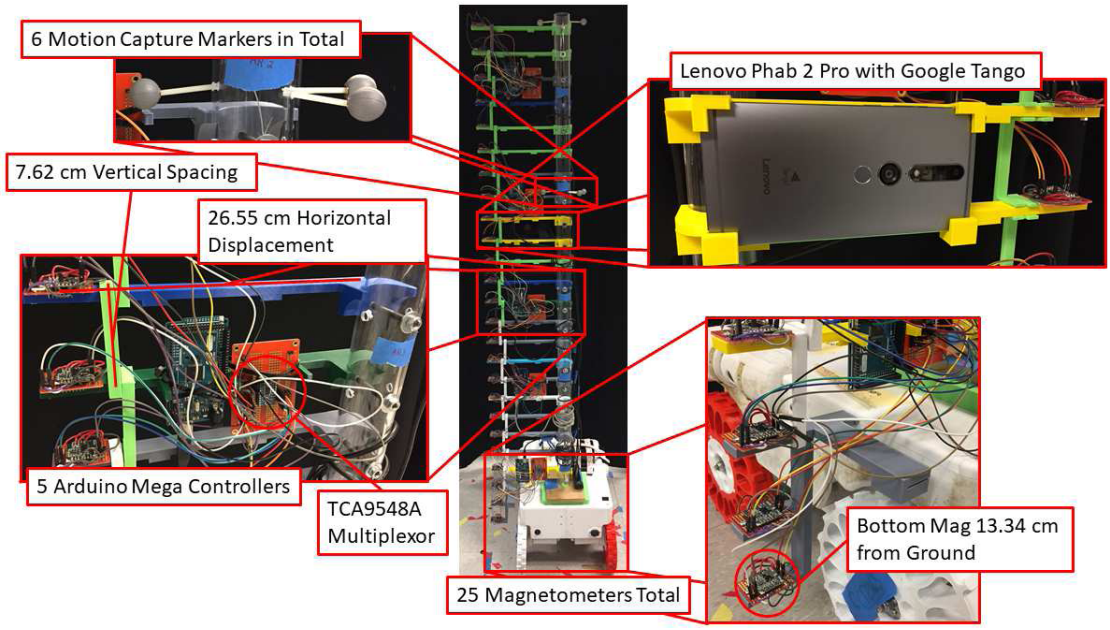


Fig. 2. TerraSentia robot along with experimental rig. A PNI RM3100 was attached to the end of each level. The robot also has a Lenovo Phab 2 Pro with Google Tango and a set of motion capture markers.

sensor at different orientations in a spatially constant magnetic field. All sensors were then aligned with the orientation of one of the magnetometers on the rig [60].

E. Data Collection

We collected data in three buildings at the University of Illinois at Urbana–Champaign: the basement floor of the CSL, the third floor of the Electrical and Computer Engineering Building (ECEB), and the third floor of Talbot Laboratory. In CSL, we surveyed one laboratory (10×9 m), four corridors (23×2 m, 22×2 m, 2×5 m, and 2×5 m), and an atrium (10×3 m), for a total of approximately 210 m^2. In ECEB, we surveyed two hallways (2×24 m and 2×24 m) and one break room (12×3 m), for a total of approximately 150 m^2. In Talbot, we surveyed five hallways (3×22 m, 1×15 m, 3×41 m, 3×18 m, and 2×11 m), for a total of approximately 290 m^2. The exteriors of these three buildings are shown in Fig. 3.

Data were collected by driving the robot to a point, stopping, and then taking many magnetometer, position, and orientation measurements at that point. Stopping the robot at each test point allowed us to further reduce confounding factors in our experimental process by eliminating error due to time synchronization and error due to flexing that may otherwise have derived from forces acting on our rig. We followed this process to collect data at 675 points in CSL, 433 points in ECEB, and 635 points in Talbot. The spatial distribution of points in each building was roughly uniform, being approximately—but not strictly—on a grid pattern. Average density, computed as the ratio of the total number of points in each building to the approximate total area surveyed in that building, was 3.2 points/m^2 in CSL, 2.2 points/m^2 in ECEB, and 2.9 points/m^2 in Talbot. The mean distance between points was 0.32 m in CSL ($\sigma = 0.13$ m, min $= 0.01$ m, and max $= 0.79$ m), 0.47 m in ECEB ($\sigma = 0.16$ m, min $= 0.02$ m, and max $= 1.07$ m), and 0.39 m in Talbot ($\sigma = 0.21$ m, min $= 0.04$ m, and max $= 1.04$ m). The mean number of sample measurements taken by each magnetometer at each point was 158 in CSL ($\sigma = 79$, min $= 17$, and max $= 1809$), 152 in ECEB ($\sigma = 31$, min $= 12$, and max $= 511$), and 325 in Talbot ($\sigma = 80$, min $= 204$, and max $= 1452$). We did *not* average these sample measurements to produce a single composite "measurement" at each point—instead, we included all sample measurements that were taken at all points in our data set.

Data were saved to a set of publicly available, comma-separated text files (IEEE DataPort, DOI: https://dx.doi.org/10.21227/t881-758110.21227/t881-7581) that were used to create magnetic maps of each building and to answer the questions posed in this article.

IV. Where in a Building Does the Magnetic Field Vary Significantly With Height?

We find that significant out-of-plane variations in the magnetic field do occur frequently and that frequency increases with the separation height of two magnetometers. The percent of points that have significantly different magnetic fields near the floor of the building is in excess of 80% of the points that we test. In fact, for most combinations of heights tested, the percent of significant test points is above 50%. Section IV-A defines what we mean by "significant" differences. Section IV-B shows that significant differences in the magnetic field occur frequently.

A. Defining "Significance"

In this section, we find that 1 μT is a significant difference in the magnetic field with respect to height. What makes a significant difference in a magnetic field depends on factors,

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

(a) Coordinated Science Laboratory (CSL). (b) Electrical and Computer Engineering Building (ECEB). (c) Talbot Laboratory.

Fig. 3. Exterior of the three buildings in which we collected data: (a) CSL, (b) ECEB, and (c) Talbot Laboratory.

TABLE I
REQUIREMENTS FOR POSITIONING (CDFXX IS PERCENT POSITIONING ERROR CUMULATIVE DISTRIBUTION FUNCTION)

Scenario	Positioning Requirement	Threshold
BVI Assistant	< 2 m at CDF100	1 μT
Retail Semantics	< 0.2 m at CDF50	1 μT

such as the localization requirements, the building, and the localization and mapping method. It is not obvious whether we should expect those significant differences to be on the order of 0.1, 1, 10 μT, and so on. The thresholds that we find here provide evidence of what can constitute a significant difference in the magnetic field. It also informs our experimental process and rig design (described in Section III). Because the threshold of significant differences played a role in experimental design, we established this threshold in prior work [35]. This section serves as a summary of this work.

The chosen threshold for significant differences in the magnetic field is motivated by two different applications of pedestrian indoor positioning technology. First, we consider the use of indoor positioning technology to serve as an assistant to the blind and visually impaired (BVI). In a recent Federal Aviation Administration report, a requirement that BVI assistants position users to within 2 m was proposed [61]. Using indoor positioning on smartphones has also been proposed to provide information to retailers regarding the effectiveness of certain displays and placement of products in stores [62]. Based on the average shelf space in U.S. convenient stores, we set the requirement for this retail positioning application to be less than 0.2 m on average [63]. A summary of scenarios and their positioning requirements is given in Table I.

We present results of having localized a pedestrian using a magnetic map produced with GPR and a particle filter. We evaluated the accuracy of the localization system by plotting the empirical cumulative distribution of the position error over a trajectory. To establish these thresholds in prior work, the GPR map is generated using data from the MagPIE data set [30]. Measurements are generated by using the same magnetic model as the map plus a constant disturbance that varies from zero to 10.0 μT. The prior uses a constant velocity motion model where the standard deviation grows at 5.6% of the distance traveled. This model is consistent with prior results presented in pedestrian dead reckoning literature [64].

On a portion of the MagPIE data set, we find that magnetic field magnitude differences above 1 μT can result in a failure to meet requirements. The empirical cumulative distribution functions (CDFs) of the positioning error that results from different disturbances are shown in Fig. 4. Therefore, in subsequent sections of this article, we will look for differences in the magnetic field on the order of 1 μT.

Note that these results may change with the environment and choice of mapping and localization algorithm. This 1-μT threshold is only an estimate of significance. Therefore, throughout the rest of this section, we will consider a threshold of 2 and 1 μT.

B. Significant Out-of-Plane Variations in the Magnetic Field Occur Frequently

Given empirical estimates of significant differences in the magnetic field, we can determine where the variation of the magnetic field is significant in two ways. First, we can use the thresholds (1 and 2 μT) to conduct a statistical t-test on the magnetic field differences with respect to height. Second, we can compute the mean and variances of the magnetic field measurements collected in each building to determine how the magnetic field varies with height. We find that, using both approaches, significant differences in the magnetic field occur frequently with respect to height.

Given that magnetic field differences on the order of a micro-Tesla can be significant in positioning applications, we perform a set of t-tests to find what percent of points in CSL, ECEB, and Talbot has these significant differences. Here, the null hypothesis is that the difference norm of the magnetic field at two heights is less than some threshold. Again, we use both 1 and 2 μT as thresholds. The details of conducting the t-tests are described in the Appendix.

Fig. 5 shows the percent of test points wherein the null hypothesis is rejected. Plots in the first row of Fig. 5 show the results as a heat map where the x- and y-axes correspond to the height of the magnetometers that are to be compared and the color is the percent of test points where the null hypothesis is rejected with a threshold of 1 μT. The same results for the CSL, the ECEB, and Talbot Laboratory are shown in the second row of Fig. 5 for a threshold of 2 μT. Here, it can be seen that the further sensors are separated with respect to height, the higher the percent of points that are significantly different.

As already mentioned, the significance thresholds that we set are dependent on a variety of factors: the choice of filter, the positioning requirements, and so on. Thus, we next consider the magnetic field without defining a significance threshold. Instead, we measure the mean and variance of

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

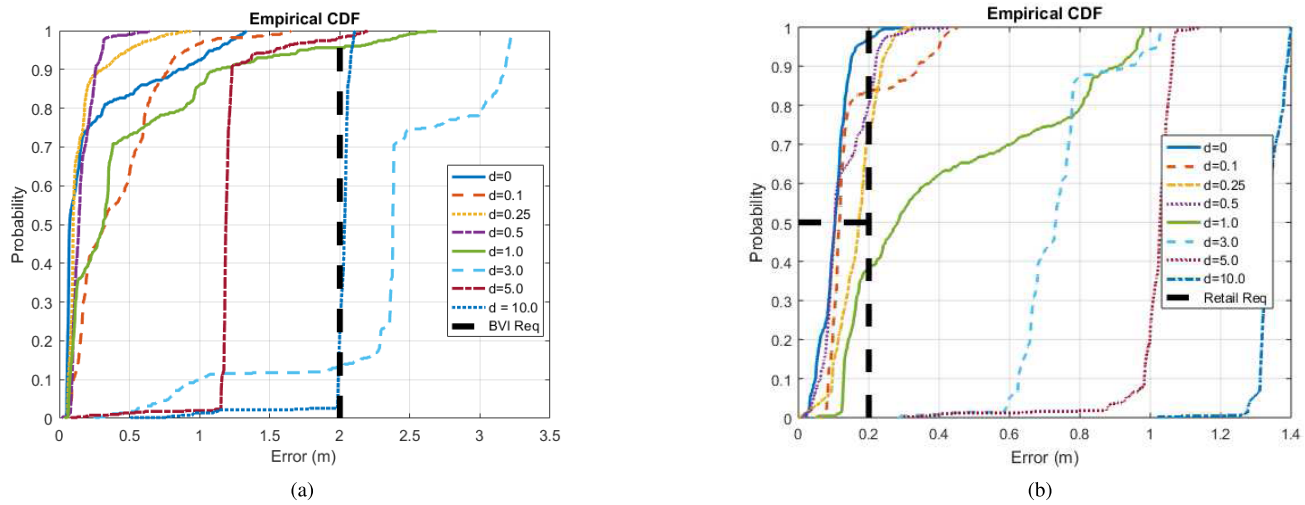


Fig. 4. Empirical cumulative distribution functions of error for a Monte Carlo localization method for different degrees of constant magnetic field magnitude disturbance. For both scenarios, we would like to ensure that this disturbance is less than 1 μT. (a) Empirical cumulative distribution of error for the BVI case, magnetic norm disturbance. (b) Empirical cumulative distribution of error for the retail case, magnetic norm disturbance.

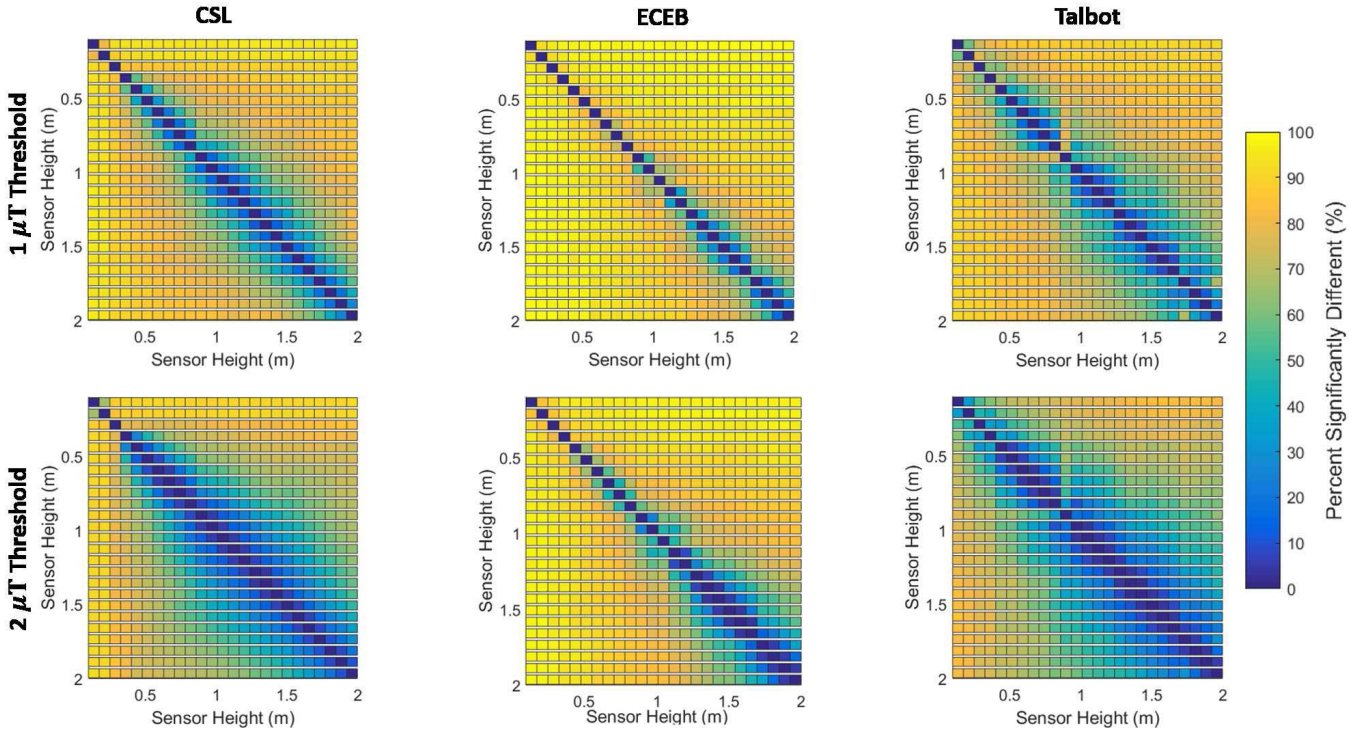


Fig. 5. Percent of test points (for each building and two different thresholds of significant differences) that have significantly different magnetic field norms between pairs of magnetometers. We can see that the further away a pair of magnetometers are from each other, the higher the percent of test points is different. Variations of the magnetic field with respect to height are also greater, the closer to the floor a magnetometer is.

the differences in the magnetic field with respect to height. Fig. 6 shows the mean difference (in absolute value) in the magnetic field at two different heights over the test points in CSL, ECEB, and Talbot. From the figure, it can be seen that the further apart two magnetometers are, the larger the magnetic difference is on average. When one magnetometer is close to the floor, the magnetic differences in these three buildings range from just over 8 μT in the Talbot Laboratory to more than 25 μT in ECEB on average. For pedestrians, the position where a smartphone is in (be it in a pocket or held in a hand) can result in a significantly different magnetic field on average. This difference could be even larger among a team of different robots or a team of pedestrians and robots working together using the same magnetic map. Fig. 6 also shows the variance in the differences over the test points in CSL, ECEB, and Talbot. From this plot, it can be seen that the variance of the magnetic field differences increases as the height difference increases. When one of the magnetometers

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

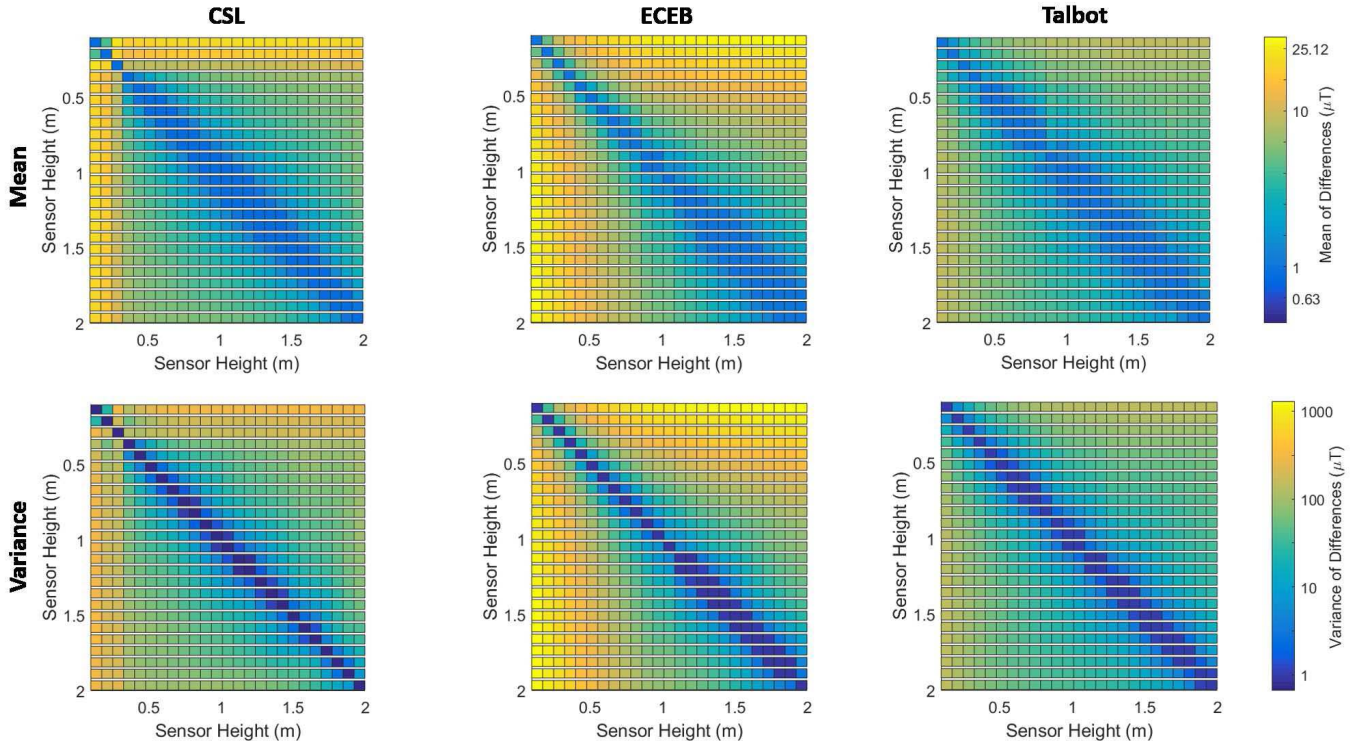


Fig. 6. Average difference (in absolute value) and variance in the magnetic field for each pair of magnetometers. Note that the difference and the variance increase, the further away the magnetometers are from each other. The color bar is set to a log scale.

is near the floor, those standard deviations can be more than 65% magnitude of the Earth's magnetic field. When one of the magnetometers is near the floor, those standard deviations can also be substantially different from one building to another: from about 11 μT in Talbot to about 35 μT in ECEB.

Note the outlier in the Talbot is plotted in Figs. 5 and 6 at a height of about 0.8 m. It is not clear why these outlier results are present; however, as previously stated, these data are available to the public for further investigation.

V. Does It Matter If the Magnetic Field Is Measured at One Height for Mapping and Another Height for Localization?

Common methods of magnetic localization use surveyed magnetic maps. We have established that the magnetic field is different at different heights. Thus, the question is not if localization accuracy will be worse if we use a map from one height and measurements for localization from a different height—it will be worse, in general. Instead, the question is how much and how quickly the localization accuracy will degrade with the difference in height.

Given our observation that the measured magnetic field changes smoothly with height (see Section IV), a reasonable hypothesis would seem to be that the ATE—a common measure of localization accuracy—will also increase smoothly with the difference in height between the magnetometer used to build a surveyed map and the magnetometer used to localize. We show in this section that this hypothesis is false. Surprisingly, for common localization methods, ATE is highly sensitive even to small differences between the map height and the localization height but stops increasing if the difference in height gets large.

We present results to support this conclusion from experiments in simulation, in which we evaluate the accuracy of localizing a ground robot under various conditions with a particle filter that is consistent with prior literature. These experiments are based on the use of magnetic maps that are constructed from the data collected in Section III. In particular, for each building, we produce 25 different magnetic maps, each one constructed by applying GPR to all magnetometer measurements taken at a given height in that building—that is to say, all measurements collected by just *one* of the 25 magnetometers on our rig. Then, given a choice of building and a simulated trajectory of the ground robot in that building, we generate a set of inertial measurement unit (IMU) measurements (often referred to as a "prior" in the literature on particle filtering) using an automotive-grade IMU model and vary two things: the height of the magnetic map that is used to *generate* simulated magnetometer measurements at a location (often referred to as a "localization model" in the literature on particle filtering), and the height of the magnetic map that is used to *evaluate the likelihood of* a given magnetometer measurement at a location (often referred to as a "map model" in the literature on particle filtering, or simply as a "map"). For each of the 625 test cases in each building—one of 25 heights for the map and one of 25 heights for the localization model—we compute the geometric mean of ATE across 20 different trajectories. Table II provides more detail about the IMU model. Table III highlights the differences between how data and models are used in this section and in subsequent sections.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

TABLE II

SENSOR MODELS USED TO PARAMETERIZE ANGLE AND VELOCITY RANDOM WALK IN OUR EVALUATIONS

Model	Price	Gyroscope Angle Random Walk	Accelerometer Velocity Random Walk
ADIS 16495 Rev. B (Tactical Grade)	\$1719	$2.61 \times 10^{-5}\ {}^{rad}/_{s\sqrt{Hz}}$	$1.336 \times 10^{-4}\ {}^{m}/_{s^2\sqrt{Hz}}$
ADIS 16334 Rev. D (Automotive Grade)	\$336	$5.80 \times 10^{-4}\ {}^{rad}/_{s\sqrt{Hz}}$	$1.837 \times 10^{-3}\ {}^{m}/_{s^2\sqrt{Hz}}$

TABLE III

SUMMARY OF THE DIFFERENT WAYS WE USE DATA AND MODELS IN MONTE CARLO LOCALIZATION AND GPR IN THE ARTICLE. HERE, h IS A HEIGHT WITH RESPECT TO THE FLOOR OF A BUILDING, h^* IS A DIFFERENT HEIGHT THAT MAY BE THE SAME OR DIFFERENT FROM h, AND Δh IS A SET OF NEIGHBORING HEIGHTS AROUND h, AS ILLUSTRATED IN FIG. 9

Section	Data used to Generate Map	Localization Model	Prior Model
Section 5	Magnetometer data at height h	Magnetometer data at height h^*	Automotive IMU
Section 6	Magnetometer data at height h	Magnetometer data at height h	Automotive and Tactical IMUs
Section 7	Magnetometer data at heights $[h - \Delta h, h + \Delta h]$	Magnetometer data at height h	Automotive IMU

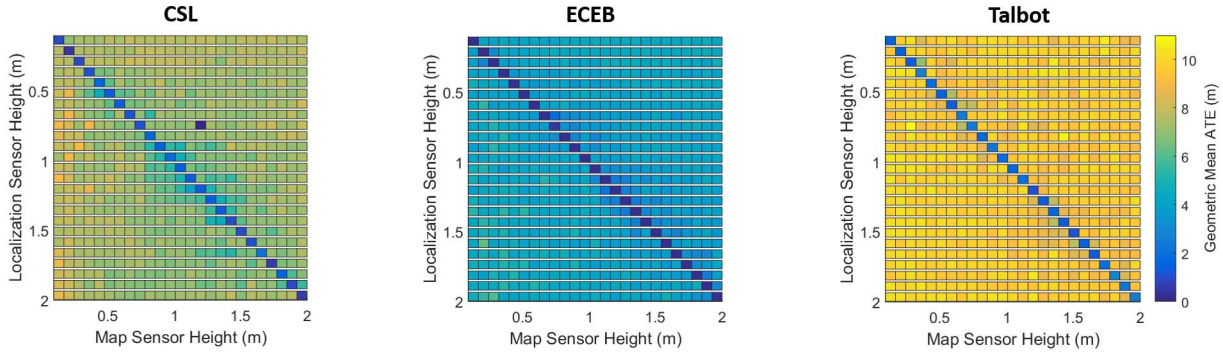


Fig. 7. Geometric mean ATE over trajectories in CSL, ECEB, and Talbot. In all cases, one axis of the plots corresponds to the height at which data were used to create a magnetic map (Map Sensor Height). The other axis of the plots corresponds to the height at which data were used to localize within that map (Loc. Sensor Height).

The Appendix provides more detail on the methods used for GPR and particle filtering (Appendix A-C), for simulating trajectories (Appendix A-E), and for computation of ATE.

Fig. 7 shows the results for each building. Since ATE cannot be negative, we assume that results at a given height follow a log-normal distribution. Again, for these common localization methods, ATE is highly sensitive even to small differences between the map height and the localization height but stops increasing if the difference in height gets large. For all three buildings, a 7.6-cm height difference between a map model and localization model (the minimum height difference used in our experiment) can have a geometric mean ATE on the order of several meters over 20 trajectories. In Section IV, we saw that, as the height between two magnetometers increases, the difference in the magnetic field at those heights also increases gradually. This trend does not extend to localization accuracy's relationship with the difference in height.

We now look more closely at several individual trials in an effort to find the cause of these unanticipated results. Fig. 8 shows a snapshot of one such trial. In particular, Fig. 8(b) compares the x-axis position of the ground-truth trajectory with the mean and standard deviation of particles. This figure shows that the mean (i.e., the state estimate) quickly deviates from the ground-truth trajectory and that the standard deviation converges and remains small—just a few centimeters—throughout. This behavior is a classic symptom of particle deprivation [65].

Given this evidence that particle deprivation is occurring, we now try to establish the reason *why*. Fig. 8(a) shows the map model and the localization model for the x-axis magnetic field as they vary along the trajectory for this same individual trial. We see immediately that the standard deviations of these two models are often less than 1 μT. There are three reasons to be alarmed by this observation. First, 1 μT is much less than the total variation of the magnetic field along the entire trajectory, which is almost 60 μT. Second, 1 μT is often less than the difference between the mean of the map model and the mean of the localization model. Third, as we saw in Section IV, differences of greater than 1 μT between magnetometer measurements at different heights are quite common in practice (see Fig. 5). We conclude that, at least in this individual trial, the map model and the localization model are simply inconsistent—most measurements will be considered statistically unlikely with respect to the map during localization. This inconsistency is exactly what we would expect to result in particle deprivation.

The standard method of solving the problem of particle deprivation, a method often referred to as *roughening*, is to add noise to resampled particle states [65, pp. 470–471]. We applied this method to the same individual trials considered above and found that it did not improve results. However,

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

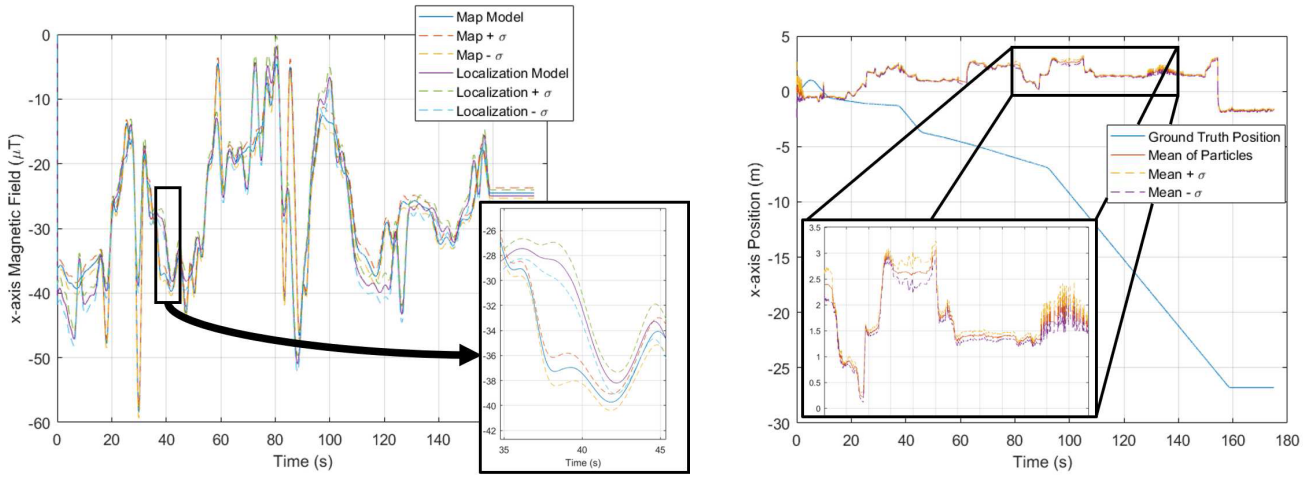


Fig. 8. Results produced by the map model, localization models, and particle filter for the second trajectory in ECEB. In this trial, the highest magnetometer on the rig is used for the map, and the fifth highest magnetometer is used for the localization model. (a) Map and localization model in the x-axis. Dashed lines represent one standard deviation above and below the mean values. The standard deviations are quite low in this case and relatively small differences in the map and the localization fall outside one standard deviation (as can be clearly seen in the inset portion of the plot). (b) Mean of the particles in the x-axis with one standard deviation of the particles above and below the mean. The ground truth is also provided. In this trial, which has a large ATE, the particles quickly collapse around a single point and diverge from the true trajectory.

TABLE IV

ATE FOR SOME INDIVIDUAL TRIALS WHEREIN ADDITIONAL VARIANCE IS ADDED TO THE MAP (MAP ROUGHEN) AND VARIANCE IS ADDED TO THE LOCALIZATION MODEL (LOCALIZATION ROUGHEN)

Building	Map Sensor	Localization Sensor	Trajectory Number	Absolute Trajectory Error (m)		
				No Roughen	Map Roughen	Localization Roughen
ECEB	1	5	2	11.29	1.42	11.74
ECEB	6	8	12	11.24	10.84	10.15
ECEB	9	15	2	11.18	4.80	11.37
CSL	16	22	5	14.87	10.97	15.90
Talbot	10	7	1	22.04	12.81	21.36

proceeding from our observation that the likely cause of particle deprivation was the inconsistency between the measurements and the map, we also tried adding noise to the map itself (by increasing its covariance) instead of to the resampled particle states and found that this variant of roughening *did* produce significantly better results. Table IV provides a summary of these results, comparing "no roughening" (the baseline case) to both "localization roughening" (the name we use here for the standard way to address particle deprivation) and "map roughening" (our heuristic variant).

These results suggest two points. First, the planar assumption can have important negative consequences on current implementations of indoor magnetic positioning suggested in the literature. Second, knowledge of the variations in the magnetic field with respect to height can be used to improve magnetic indoor positioning algorithms.

We emphasize that the heuristic method of "map roughening" that we proposed here is only one possible way of ameliorating the impact of height on an indoor magnetic positioning system and that our preliminary results do not allow us to conclude that this method would be helpful in all cases. There are many opportunities to explore the application of both this and other methods—for example, localization with a particle filter that incorporated a covariance estimation strategy [66]—in future work, using our results as a baseline for comparison.

The results that we presented in this section were unanticipated, in part, because a large sensitivity of ATE with respect to small changes in height has not been observed in prior work. We do not know exactly why this sensitivity has not been observed previously. However, our results suggest that the magnetic field surveys used in prior work may have generated maps with higher covariance than the map that we generated with our own magnetic field survey. For example, some of these prior surveys were conducted by pedestrians with hand-held devices, who were asked to walk through buildings without stopping. Data produced by such a survey are likely to be sparser—with fewer measurements near any given point—and noisier (e.g., due to natural variations in height while walking) than data produced by our survey. The effect may be to create "roughened" magnetic maps, which we have shown can significantly reduce the sensitivity of ATE to changes in height when used for localization. Assuming that our reasoning is correct, the problem here is not with our survey, which more accurately measures the magnetic field than surveys used in most prior work. Rather, the problem is that standard methods of localization produce the unfortunate

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

TABLE V
SPEARMAN CORRELATION COEFFICIENT AND p-VALUES OF THE GEOMETRIC MEAN OF THE ATE AS A FUNCTION OF HEIGHT FOR EACH OF OUR SIX TEST CASES

Test Case	Spearman Coefficient	p-value
Auto IMU, CSL	-0.0254	0.9048
Auto IMU, ECEB	-0.1300	0.5341
Auto IMU, Talbot	0.0838	0.6894
Tactical IMU, CSL	0.2215	0.2858
Tactical IMU, ECEB	-0.3738	0.0664
Tactical IMU, Talbot	-0.4800	0.0162

result that, as a map gets better, localization accuracy can get much worse. Again, further study is warranted.

VI. DOES IT MATTER AT WHICH HEIGHT THE MAGNETIC FIELD IS MEASURED, IF THIS HEIGHT IS KEPT THE SAME FOR BOTH MAPPING AND LOCALIZATION?

In robot design, the location of a magnetometer can be chosen. If the magnetic field varies with height; then, it may be the case that one height may be better for localization than other heights. Here, we define "better" height as the height that tends to produce the most accurate location estimate using methods common in other studies. It is reasonable to assume that the ATE—our measure of accuracy—computed using measurements from magnetometers near the floor of a building would be lower than the ATE computed using measurements from magnetometers higher on our rig. This hypothesis is due to the fact that we know that many structural members are in the floors of buildings, and the results in Fig. 6 could be produced by stronger magnetic fields in the floor of buildings. We show that, for the three case studies, the location of the magnetometer (with respect to height) does not have a significant impact on ATE. Rather, the choice of the sensor (such as a tactical versus an automotive-grade IMU) or the choice of the building has a much more pronounced effect.

As with the prior section, we present results of having localized a ground robot using a magnetic map produced with GPR and a particle filter that is consistent with prior literature. In this section, the GPR map is generated using data from each height (for a total of 25 different maps). The localization model is generated using the same GPR model as the map. We use two different sensors in our prior: a tactical and automotive-grade IMU, as described in Table II. For each building, each height, and each trajectory, ATE is computed. Therefore, there are 3000 tests. We evaluate the accuracy of the localization system by computing ATE over 20 randomly generated trajectories. Again, we describe these methods in detail in the Appendix. Table III highlights the differences between how data and models are used in this section as opposed to prior sections.

We compute the geometric mean of the ATE across all 20 test trajectories. We use a geometric mean because ATE cannot be negative; therefore, we assume that the distribution of results at a given height follows a log-normal distribution. We compute the Spearman correlation coefficient of ATE as a function of height and their associated p-values for all six test cases. These results are listed in Table V. Note that these results show weak correlations for five of the six test cases. The p-values for these five test cases are also above 0.05. Only for the tactical IMU in Talbot, there is a moderate correlation of -0.48 and a p-value (0.0162) below a threshold of 0.05. Applying linear regression to the tactical IMU in the Talbot case, the slope of this trend is -0.18 cm of ATE per centimeter of height. These trends, or lack thereof, refute our original hypothesis.

While these results do not confirm our initial belief, other results are as expected. From all three plots, it can be observed that, with respect to the geometric mean, a tactical IMU produces a lower ATE in all three buildings than an automotive-grade IMU. In addition, the geometric mean ATE in ECEB is smaller than in the other two buildings. We would expect there to be noticeable differences between buildings since these differences have been apparent in prior literature. Also, unsurprisingly, the magnetic map adds significant value to the navigation system—using the IMUs alone (in other words, without resampling), the geometric mean ATE in all three buildings is substantially higher than with the magnetometers.

We do not know what was special about the single test case—using a tactical IMU in Talbot—for which there was a statistically significant correlation between height and ATE. These results certainly do confirm that the choice of height at which the magnetic field is measured *can* matter significantly in some environments and for some sensors. Future work might help to resolve the open question about exactly which environments and which sensors by considering a larger number of buildings—particularly buildings of different types (e.g., warehouses, homes, or high-rises)—and a wider variety of IMU grades.

Note that variations in the magnetic field due to ferromagnetic objects in the floor likely only have a local effect on the ability to localize, so the choice of the trajectory may be very important to leveraging greater magnetic variations that exist near the floor of a building. Thus, for example, if one were to generate a random path between two points (as we do here), there may not be any relationship between ATE and height on average. However, if we were to plan a path between two points such that the robot crosses near or over the greatest number of magnetic anomalies, there may be a noticeable difference between ATE at different heights. If this hypothesis were true, it would be possible that a few randomly generated trajectories would have statistically significant trends with respect to height, which is apparent in our results. For example, what we label the 13th trajectory in CSL has a very strong correlation (the Spearman coefficient of 0.82) when a tactical IMU is used. This correlation is statistically significant, with a p-value of 2.3×10^{-6}. Applying a linear regression in this case, the slope, in this case, is 0.20 cm of ATE per centimeter of height.

There may also be a noticeable relationship between height and other metrics besides ATE. For example, the circular error 95% (CE95) may be related to a height even though the geometric mean ATE is not. Other important localization metrics are identified in the International Organization for

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

Standardization (ISO)/ International Electrotechnical Commission (IEC) 18305, and path-dependent metrics are presented by Mendoza-Silva *et al.* [67]. Finally, we consider the primary error source to be accelerometer velocity random walk and gyroscope angle random walk in this section. There are other possible errors that may have been chosen and may also have a relationship with height. Accelerometer or gyroscope bias or rate random walk may also be considered for example. Sensor bias is often removed with calibration prior to use or is estimated online. Rate random walk is not universally considered in magnetometer-based localization. Sensor noise, on the other hand, is universally considered in all magnetometer-based localization methods, which is why we vary these parameters.

VII. Does It Matter If Measurements of the Magnetic Field at Different Heights Are Aggregated to Produce a Single, Planar Map?

In this section, we consider the case where data collected at multiple, neighboring heights are used to generate a magnetic map. This case is relevant for crowd-sourcing maps from multiple pedestrians [23]. By crowd-sourcing maps, we mean using magnetometer and IMU data collected from multiple smartphones used by multiple individuals to generate a map of the magnetic field of a building in the cloud. This question is also relevant if different robots are used to generate a magnetic map or if teams of pedestrians and robots are used to generate and localize on magnetic maps. In particular, for a magnetometer at a given height, h, from the floor, we use magnetometer data from sensors within the range $[h - \Delta h, h + \Delta h]$ to generate a map. We call $\Delta h > 0$ the neighboring height. Fig. 9 illustrates the scenario we consider here.

If we take the planar assumption to be false, as shown in Section IV, then it would be reasonable to suspect that, as sensors at different heights are aggregated together, localization error would increase. However, the results presented in Sections V and VI provide reasons to propose a different set of hypotheses. In Section V, we argued that adding covariance to surveyed map models may produce better localization results. Therefore, we hypothesize here that, as data from neighboring heights are used to generate a map, our measure of localization accuracy, ATE, decreases. We hypothesize this result because data from neighboring heights will likely increase the covariance of the map. Furthermore, we hypothesize that, as the number of neighboring magnetometers used to generate a map increase, ATE eventually increases again. More specifically, we propose that ATE as a function of Δh (and at a given height h) follows a convex quadratic function. Moreover, we also considered the height, h, at which mapping a magnetic field and localizing a robot on that surveyed map produces the lowest ATE in Section VI. We found height does not have a significant impact on ATE. Because ATE is not significantly impacted by height, we hypothesize further that ATE as a function of Δh is a convex quadratic function with respect to all heights in a building (independent of the choice of h). In other words, if we find that ATE as a function of Δh for the sixth magnetometer from the floor follows a quadratic function, that function also reasonably approximates the relationship between ATE and Δh at say the twentieth magnetometer from the floor. We show in this section that these hypotheses are true for these three buildings. First, we will show that ATE as a function of Δh follows a convex quadratic function for a given height h. Then, we will show that ATE as a function of Δh follows a convex quadratic function independent of a given height h.

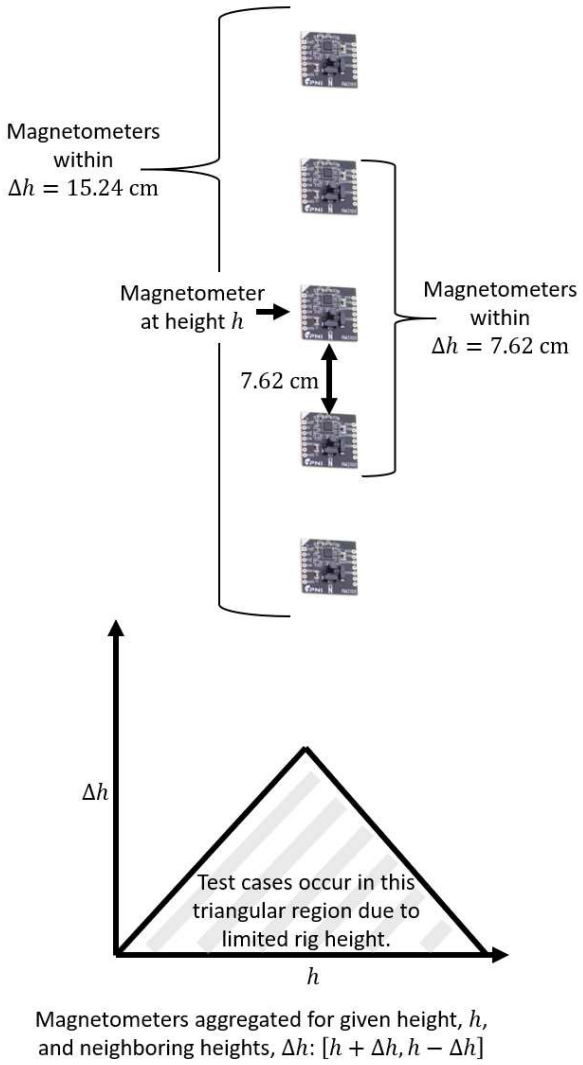
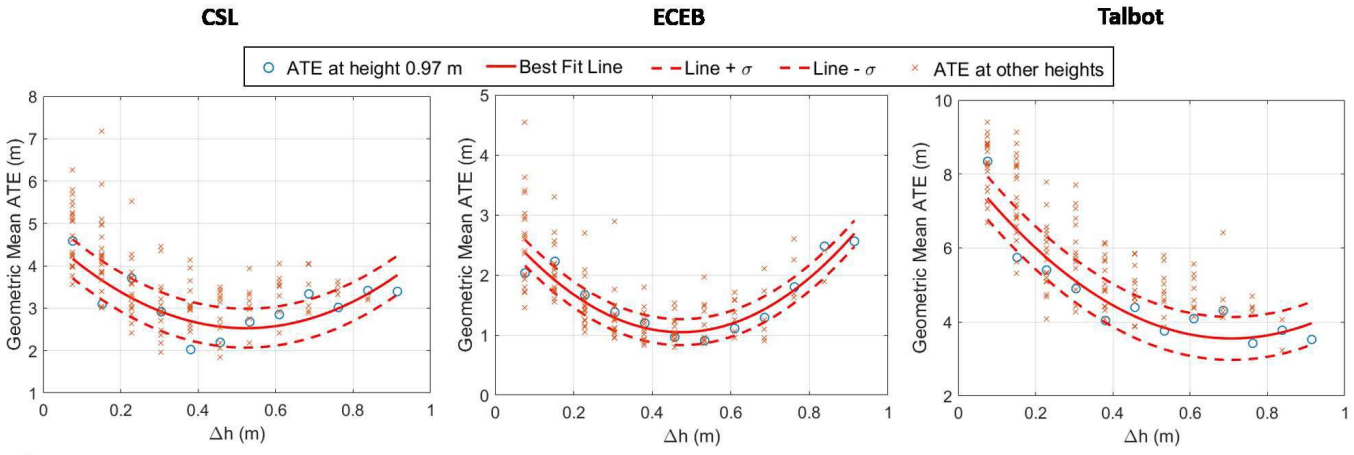


Fig. 9. Data in Section VII are aggregated by first selecting a sensor a given height, h, from the floor. Then, a neighboring height Δh is chosen, and the magnetometer data from sensors within the range $[h - \Delta h, h + \Delta h]$ are used for map generation.

In summary, we observed in Section V that map roughening (i.e., adding covariance to surveyed magnetic maps) improved ATE relative to a baseline particle filter. We hypothesize that aggregating sensor measurements at nearby heights effectively "roughens" the surveyed magnetic map, thereby improving ATE. However, we believe that this is only true up to a point. If magnetometers from large height differences are aggregated, this would increase the covariance of the magnetic map too much and begin to degrade the ATE.

As with prior sections, we present results of having localized a ground robot using a magnetic map produced with GPR

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.



Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	3.02	1	3.02	14.37	0.00428
x	3.19	1	3.19	15.20	0.00363
Error	1.89	9	0.21		
Total	8.10	11			

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	3.19	1	3.19	68.41	1.69×10^{-5}
x	2.76	1	2.76	59.09	3.04×10^{-5}
Error	0.42	9	0.047		
Total	6.37	11			

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	4.11	1	4.11	12.20	0.00679
x	7.92	1	7.92	23.54	0.000906
Error	3.03	9	0.34		
Total	15.06	11			

Fig. 10. ATE at different heights as a function of the height of neighboring magnetometers Δh used to generate a magnetic map. A best fit line produced using ATE gathered at a height of 0.97 m is also presented. Points used for regression are presented as circles. In CSL, $R^2 = 0.63$. In ECEB, $R^2 = 0.89$. In Talbot, $R^2 = 0.85$.

TABLE VI

ANALYSIS OF VARIANCE OF SECOND-ORDER POLYNOMIAL REGRESSION IN CSL USING DATA AT ALL HEIGHTS, h. $R^2 = 0.46$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	27.73	1	27.73	57.59	4.00×10^{-12}
x	42.36	1	42.36	87.95	1.54×10^{-16}
Error	67.90	141	0.48		
Total	137.99	143			

and a particle filter that is consistent with prior literature. In this section, the GPR map is generated using data from each height, h, plus magnetometers located up to a height Δh immediately above and below a given magnetometer. The localization model is generated using data at height h. We use an automotive IMU as our prior. For each building, each height h, each set of neighboring heights Δh that exist, and each trajectory, ATE is computed. Therefore, there are 8640 tests. We evaluate the accuracy of the localization system by computing ATE over 20 randomly generated trajectories. Again, we describe these methods in detail in the Appendix. Table III highlights that the differences between how data and models are used in this section as opposed to prior sections.

We compute the geometric mean of the ATE across all 20 test trajectories and plot them in Fig. 10. To test our hypotheses, we perform regression on the heights with the largest number of neighboring magnetometers (which happens to be 0.97 m from the floor). We assess the validity of our first hypothesis by regressing a second-order polynomial to our data, and we compute the coefficient of determination for the results in all three buildings and analysis of variance. These results are also presented in Fig. 10. In all three buildings, a second-order polynomial accounts for a substantial majority of the variability seen in the data (as measured by the coefficients of determination). Moreover, the computed p-values in the analysis of variance justify accepting the proposed second-order polynomial model predicted by our hypothesis.

To evaluate if ATE as a function of Δh follows a convex quadratic function independent of a given height h, we expand the data we use for regression to ATE collected at every height h. In other words, we regress on all points marked "x" and "o" in Fig. 10 and evaluate if there is still a statistically significant quadratic trend. The result of an analysis of variance of the regression is shown in Tables VI–VIII. Unsurprisingly, the coefficient of determination decreases for all three cases. However, the p-values for both terms in the quadratic function also decrease substantially. These results support the claim that, across all heights h, ATE follows a convex quadratic function when magnetometer data are aggregated over neighboring height Δh.

The results presented in this section demonstrate that localization accuracy improves if we aggregate data collected at different heights into a single, 2-D map: up to a point. If the range of heights aggregated is too large, the localization error may increase. The optimal height appears to vary from building to building based on our regressed functions. Aggregation is just one example of how the planar assumption can be used in localization. It requires the deliberate collection of magnetometers at a range of heights. Thus, while the planar assumption may be strictly speaking false, this does not mean

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

TABLE VII
ANALYSIS OF VARIANCE OF SECOND-ORDER POLYNOMIAL REGRESSION IN ECEB USING DATA AT ALL HEIGHTS, $h.R^2 = 0.57$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	27.53	1	27.53	134.25	3.15×10^{-22}
x	34.97	1	34.97	170.51	4.84×10^{-26}
Error	28.92	141	0.21		
Total	92.42	143			

TABLE VIII
ANALYSIS OF VARIANCE OF SECOND-ORDER POLYNOMIAL REGRESSION IN TALBOT USING DATA AT ALL HEIGHTS, $h.R^2 = 0.70$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	p-Value
x^2	33.83	1	33.83	44.48	5.40×10^{-10}
x	91.04	1	91.04	119.71	1.49×10^{-20}
Error	107.24	141	0.76		
Total	232.11	143			

that a planar map should not be used. Rather, the implications of making a planar assumption are nuanced. One additional nuance that relates to these results is that the conclusions reached here are likely dependent on the choice of filter. This dependence is why we choose a filter that is consistent with prior literature. However, alternative filters may produce very different results.

VIII. CONCLUSION

We provide detailed, quantitative characterizations regarding the implications of using a planar assumption in magnetic field-based localization in this article. We find that significant variations in the magnetic field occur with respect to height changes in all three buildings. When we generate a magnetic map at one height and localize at a second height, ATE is sensitive even to small differences in heights. We also find that the choice of height at which to take measurements—if this height was kept the same for both localization and mapping—had no significant impact on ATE when averaged across a given set of trajectories although some trajectories existed for which different measurement heights led to significantly different errors. Finally, we find that ATE decreased when magnetometer measurements were aggregated across a small range of heights to produce a single, planar map and when measurements at the median height were used for localization.

One of the implications of this work is how it may inform the testing and evaluation of magnetic localization systems. There exists an international standard for the test and evaluation of indoor positioning systems [33]. In this standard, it is said that "it is important to: 1) identify the circumstances under which a location sensor does not work well or outright fails and 2) ensure the localization and tracking system test and evaluation scenarios include vignettes where those failures are likely to happen." We have shown that when evaluating the performance of a magnetic localization system, especially one using the planar assumption, it is important to consider how the height of the magnetometer changes during operation. Data sets used to evaluate magnetic localization systems should include changes in height that are anticipated during the operation of a device.

In terms of the methods using magnetic localization, the descriptions in this work suggest some further hypotheses that can be evaluated in future work.

- *H1:* Magnetic SLAM methods that use gesture recognition to identify whether a phone is held in one's hand, to one's ear, or in one's pocket and switch between three different magnetic maps can improve the ATE relative to methods that create one, planar magnetic map.
- *H2:* Using a 3-D magnetic map to localize a user can improve the ATE relative to methods using one, planar magnetic map.
- *H3:* Differences in the magnetic field with respect to height may be clustered. If differences in a magnetic field with respect to height appear to be clustered, a Monte Carlo localization strategy with outlier rejection on magnetometer measurements may decrease ATE. Some examples of these outlier rejecting localization strategies are described in prior work [68]–[70].

In this work, we do not investigate what materials, configurations of materials, and other properties of a building cause the observations we make. Understanding the root causes behind our observations could allow us to predict the kind of magnetic field variations that we can expect in buildings. Such predictions would allow us to improve test and evaluation methods, use accurate and scientifically grounded assumptions in localization methods, and predict the limitations of this technology. One way to make such causal links would be to investigate using a building information model to predict the magnetic field in a building. This approach would allow us to associate magnetic field variations directly with certain building elements, such as beams, HVAC components, and rebar. Another means by which the study of magnetic field variations could be extended would be to conduct the study in this article in different types of buildings. The ISO/IEC 18305:2016(E) standard recommends testing in five different types of buildings: a single-family house, a three-story brick and concrete office building, a warehouse or factory, a high-rise steel structure, and a subterranean structure. It is possible that the results that we present here may differ in buildings in these different categories. Alternatively, it may be the case that these categories are not meaningful in the

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

TABLE IX
SUMMARY OF CONFOUNDING FACTORS

Confounding Factor	Description of Impact	Magnitude Estimate
Magnetometer Hysteresis	Hard iron bias will change during test	0.00375%
Magnetometer Bias Random Walk	Hard iron bias will change during test	$< 0.008^{\mu T}/_{s\sqrt{Hz}}$
Ground Truth Positioning Error	Incorrect locations of magnetic measurements	$\sim 5.9\ cm$ RMS ATE
Earth Magnetic Model Error	Magnetometer measurement scale factor error	$< 0.09\ \mu T$

context that we consider in this article. Future work could identify categories of buildings that allow designers to predict the degree to which the magnetic field may vary with height.

We acknowledge that all questions posted in this article were answered with respect to a single choice of magnetometer—the PNI RM3100—as described in Section III-A. This magnetometer, while an appropriate choice for many ground robots and for some hand-held devices, was chosen because of its high accuracy—in particular, because of its lower noise and higher sensitivity—relative to magnetometers used in prior work. It is natural to ask how (if at all) our results would have changed had we used a magnetometer with lower accuracy. Since it would be straightforward to simulate the use of a lower-accuracy magnetometer by artificially adding noise to our publicly available data set, a systematic study of how results vary with accuracy seems like one more good opportunity for future work.

APPENDIX A
SUPPLEMENTAL INFORMATION

A. Summary of Confounding Factors

Confounding factors are identified and discussed throughout the text. In Table IX, those factors are collected and summarized. The magnitudes in the table quantify the confounding factors themselves and not the impact it has on the experiment. For example, the magnetometer hysteresis is quantified as a percent of the applied magnetic field rather than in terms of the resulting hard iron bias change that may occur during a test.

B. Additional RM3100 Sensor Details

Performance specifications for our sensors as reported by PNI Corporation are listed in Table X. We also recorded measurements from all 25 magnetometers for 1 h without moving the sensors or any possible magnetic source near the sensors. We plotted the Allan deviation of these sensors in Fig. 11. The Allan deviation (much like power spectral density) is a common description of sensor noise properties. From these plots, it is clear that so-called rms noise and rate random walk are slightly different from one sensor to the next. The presence of rate random walk noise in our sensor is a notable confounding factor.

C. Details on Localization and Mapping

To produce the localization results in this article, we apply a Monte Carlo localization method with 4000 particles and low variance resampling, as described by Thrun *et al.* [72], and auxiliary weights, as described by Simon [65]. The

TABLE X
PNI RM3100 REPORTED VALUES [71]

Parameter	Value	Units
Field Measurement Range	± 800	μT
Sensitivity	0.013	μT
Noise	0.015	μT
Repeatability (over $\pm 200 \mu T$)	0.008	μT
Hysteresis (over $\pm 200 \mu T$)	0.015	μT
Operating Temperature	-40 to 85	$^\circ C$

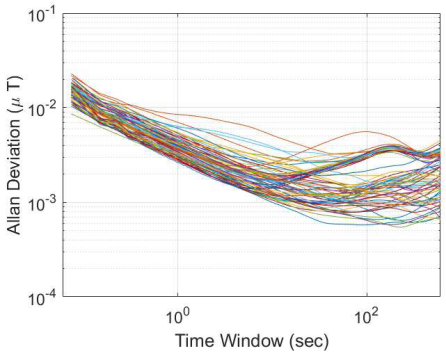


Fig. 11. Allan deviation plots for our 25 magnetometers show that our sensors have different rms noises and rate random walks.

a priori sampling was performed using a modeled accelerometer, gyroscope, and 2-D, nonslipping car. This approach is consistent with Monte Carlo approaches used in prior literature on magnetic-field-based localization, as described in Table XI. Mapping is performed via GPR using the package GPy [73]. Since there is a large amount of magnetometer data, the stochastic variational inference is used in GPy to learn a map with a radial basis function, automatic relevance determination, and a constant bias [74]. Note that, as described in Section III-E, this approach was applied to all sample measurements collected at all points in each building (rather than to measurements averaged pointwise, for example). This approach was performed for all three axes of the building's frame of reference independently.

D. Details on Statistical t-Test

We use a nonstandard t-test in that we test the hypothesis that the difference between two population means lay within a range of values (whereas t-tests typically test the hypothesis that the differences between the means of two populations are equal). Because our approach is nonstandard, we cover the details of our t-test in significant detail and highlight where our approach differs from a standard approach. We begin with assumptions for two-sample inference.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

TABLE XI

MONTE CARLO LOCALIZATION USED IN PRIOR LITERATURE AND THE APPROACH USED IN THIS ARTICLE. THE APPROACH USED HERE IS CONSISTENT WITH THE LISTED METHODS

	This Paper	**Haverinen and Kemppainen [9]**	**Wang et al [12]**	**Vallivaara et al [17]**	**Frassl et al [75]**
Prior	IMU dead reckoning	Prior state plus samples from normal distribution	unknown	unknown	Wheel encoder odometry
Number of Particles (N_p)	4000	4000	1000	$<$1000	unknown
Effective Particle threshold	$N_p/2$	$N_p/2$	unknown	unknown	$N_p/2$
Resampling	Low var. resampling with auxiliary weights	unknown	unknown	unknown	systematic
Map	Gaussian Process Regression	Gaussian dist. plus linear interpolation	Gaussian Process Regression	Gaussian Process Regression	Gaussian dist. plus linear interpolation

1) $X_{11}, \ldots, X_{1n_1}$ is a random sample from population 1.
2) $X_{21}, \ldots, X_{2n_2}$ is a random sample from population 2.
3) The two populations represented by X_1 and X_2 are independent.
4) Both populations are normally distributed.

We say that $\bar{X}_1$ and $\bar{X}_2$ are the sample means of population 1 and population 2, respectively. The sample variances of population 1 and population 2 are S_1^2 and S_2^2 respectively. The expected value of the difference in sample means and the variance of the differences are

$$\mathbb{E}\big(\bar{X}_1 - \bar{X}_2\big) = \mu_1 - \mu_2$$
$$\mathrm{Var}\big(\bar{X}_1 - \bar{X}_2\big) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Converting the distribution of the differences in sample means to a standard normal distribution

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathcal{N}(0, 1).$$

Everything above is typical in t-tests (see [76]). We wish to test the null hypothesis

$$H_0 : \mu_1 - \mu_2 = \Delta \in \big[\mu_{1,f} - \mu_{2,f} - \delta, \mu_{1,f} - \mu_{2,f} + \delta\big]$$

where $\mu_{1,f}$ and $\mu_{2,f}$ are the sample means of the magnetometer measurements in the constant magnetic field after calibration, which is not standard. Plugging in for the expected value of the difference in sample means $(\mu_1 - \mu_2)$, we get

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Using the sample variances, we arrive at a t-statistic

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}.$$

This t-statistic is the same as that derived in Montgomery and Runger [76]. However, unlike Montgomery and Runger, because Δ is a range of values, t_0 is also a range of values [76]. We propose to use a conservative hypothesis test. We reject the null hypothesis only if there exists no choice of Δ such that a standard two-sided t-test would accept the null hypothesis. In other words, we reject the null hypothesis if

$$t_0 = \frac{\bar{X}_1 - \bar{X}_2 - \mu_{1,f} + \mu_{2,f}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

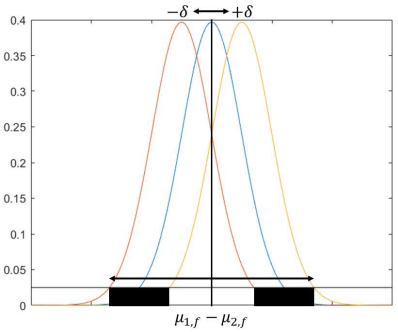


Fig. 12. Our null hypothesis states that differences in the means of two populations belong to a range $[\mu_{1,f} - \mu_{2,f} - \delta, \mu_{1,f} - \mu_{2,f} + \delta]$. Drawing a vertical black line at significance level 0.05 for each distribution produces the black boxes in the figure. We reject the null hypothesis if the t-value is to the left or right of both boxes.

and

$$t_0 > t_{\frac{\alpha}{2},\nu} + \frac{\delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \text{ or}$$
$$t_0 < -t_{\frac{\alpha}{2},\nu} - \frac{\delta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}.$$

This test is conservative in the sense that we could assume that Δ could be a random variable with some probability distribution (say a uniform distribution). A t-test derived from such an assumption would reject the null hypothesis for smaller values (in an absolute sense) of t_0. The intuition of our t-test is shown in Fig. 12. Note that an alternative interpretation of our hypothesis test is to assume that the null hypothesis is

$$H_0 : \mu_1 - \mu_2 = \mu_{1,f} - \mu_{2,f}$$

and to choose a significance level α according to the previously described criterion.

E. Generation of Test Trajectories

We test our ability to localize by generating 20 different sample trajectories each for CSL, ECEB, and Talbot. Each

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

of our trajectories is generated by tracking a path between two hall doors, entrances, or other natural starting points. We assume that the dynamics of a ground robot can be described by the Dubin's car model

$$\begin{aligned}\dot{x} &= v\cos\theta \\ \dot{y} &= v\sin\theta \\ \dot{v} &= u_1 \\ \dot{\theta} &= \frac{v\tan u_2}{L}.\end{aligned}$$

We assume that the range of steering in this vehicle is

$$-\frac{\pi}{2} + 0.01 \leq u_2 \leq \frac{\pi}{2} - 0.01$$

and that the maximum turning speed is 0.6 rad/s. The desired forward velocity is 0.3 m/s and is controlled using the basic law

$$u_1 = k(v_{\text{des}} - v)$$

where $k = 10$. We use rapidly exploring random trees (RRT) to construct a viable path from each endpoint.

F. Absolute Trajectory Error

We use the ATE as the standard measurement of localization error throughout this article. This is a commonly used metric for localization error, as described in detail by Sturm *et al.* [57]. Given an estimated trajectory $\mathbf{P}_{1,\ldots,n}$ for n time steps, $\mathbf{P} \in SE(2)$, ground-truth trajectory $\mathbf{Q}_{1,\ldots,n}$, and $\mathbf{Q} \in SE(2)$, we first apply Horn's method to find the transformation $\mathbf{S} \in SE(2)$ that minimizes the least-squares error between $\mathbf{P}_{1,\ldots,n}$ and $\mathbf{Q}_{1,\ldots,n}$. The ATE at time step i is

$$\text{ATE}_i = \mathbf{Q}_i^{-1}\mathbf{S}\mathbf{P}_i.$$

The ATE over an entire trajectory is the root mean square error of the translational components of the ATE over all time steps in the trajectory

$$\text{ATE} = \left(\frac{1}{n}\sum_{i=1}^{n}\|\text{trans}(\text{ATE}_i)\|^2\right)^{1/2}.$$

ACKNOWLEDGMENT

The authors would like to thank Scott D. Zelman and Siddharth Chadha for help constructing portions of our experimental rig. This work made use of the Illinois Campus Cluster, a computing resource that is operated by the Illinois Campus Cluster Program (ICCP) in conjunction with the National Center for Supercomputing Applications (NCSA) and that is supported by funds from the University of Illinois.

REFERENCES

[1] S. He and K. G. Shin, "Geomagnetism for smartphone-based indoor localization: Challenges, advances, and comparisons," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 97:1–97:37, Dec. 2017.

[2] 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, pp. 651–667, Mar. 2017.

[3] A. Solin, M. Kok, N. Wahlström, T. B. Schön, and S. Särkkä, "Modeling and interpolation of the ambient magnetic field by Gaussian processes," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1112–1127, Aug. 2018.

[4] S. A. Rahok, Y. Shikanai, and K. Ozaki, "Navigation using an environmental magnetic field for outdoor autonomous mobile robots," *Adv. Robot.*, vol. 25, nos. 13–14, pp. 1751–1771, Jan. 2011.

[5] S. A. Rahok, K. Inoue, and K. Ozaki, "Development of a mobile robot to run in tsukuba challenge 2010," *Adv. Robot.*, vol. 26, no. 14, pp. 1555–1575, Sep. 2012.

[6] K. P. Subbu, B. Gozick, and R. Dantu, "LocateMe: Magnetic-fields-based indoor localization using smartphones," *ACM Trans. Intell. Sys. Technol.*, vol. 4, no. 4, pp. 73:1–73:27, Oct. 2013.

[7] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, and M. Wiseman, "Indoor location sensing using geo-magnetism," in *Proc. 9th Int. Conf. Mobile Syst., Appl., Services (MobiSys)*, 2011, pp. 141–154.

[8] C. Gao and R. Harle, "MSGD: Scalable back-end for indoor magnetic field-based GraphSLAM," in *Proc. Int. Conf. Robot. Autom.*, May 2017, pp. 3855–3862.

[9] J. Haverinen and A. Kemppainen, "Global indoor self-localization based on the ambient magnetic field," *Robot. Auto. Syst.*, vol. 57, no. 10, pp. 1028–1035, Oct. 2009.

[10] E. Le Grand and S. Thrun, "3-axis magnetic field mapping and fusion for indoor localization," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Sep. 2012, pp. 358–364.

[11] B. Kim and S.-H. Kong, "A novel indoor positioning technique using magnetic fingerprint difference," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 2035–2045, Sep. 2016.

[12] X. Wang, C. Zhang, F. Liu, Y. Dong, and X. Xu, "Exponentially weighted particle filter for simultaneous localization and mapping based on magnetic field measurements," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1658–1667, Jul. 2017.

[13] N. Lee and D. Han, "Magnetic indoor positioning system using deep neural network," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2017, pp. 1–8.

[14] R. Montoliu, J. Torres-Sospedra, and O. Belmonte, "Magnetic field based indoor positioning using the bag of words paradigm," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2016, pp. 1–7.

[15] D. Carrillo, V. Moreno, B. Úbeda, and A. Skarmeta, "MagicFinger: 3D magnetic fingerprints for indoor location," *Sensors*, vol. 15, no. 7, pp. 17168–17194, Jul. 2015.

[16] S. Wang, H. Wen, R. Clark, and N. Trigoni, "Keyframe based large-scale indoor localisation using geomagnetic field and motion pattern," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 1910–1917.

[17] I. Vallivaara, J. Haverinen, A. Kemppainen, and J. Röning, "Magnetic field-based SLAM method for solving the localization problem in mobile robot floor-cleaning task," in *Proc. 15th Int. Conf. Adv. Robot. (ICAR)*, Jun. 2011, pp. 198–203.

[18] S. Shahidi and S. Valaee, "GIPSy: Geomagnetic indoor positioning system for smartphones," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2015, pp. 1–7.

[19] S.-E. Kim, Y. Kim, J. Yoon, and E. Sun Kim, "Indoor positioning system using geomagnetic anomalies for smartphones," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Nov. 2012, pp. 1–5.

[20] J. Jung, S.-M. Lee, and H. Myung, "Indoor mobile robot localization and mapping based on ambient magnetic fields and aiding radio sources," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 1922–1934, Jul. 2015.

[21] N. Akai and K. Ozaki, "Gaussian processes for magnetic map-based localization in large-scale indoor environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2015, pp. 4459–4464.

[22] P. Robertson *et al.*, "Simultaneous localization and mapping for pedestrians using distortions of the local magnetic field intensity in large indoor environments," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2013, pp. 1–10.

[23] A. Ayanoglu, D. M. Schneider, and B. Eitel, "Crowdsourcing-based magnetic map generation for indoor localization," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Nantes, France, Sep. 2018, pp. 1–8.

[24] F. Al-Homayani and M. Mahoor, "Improved indoor geomagnetic field fingerprinting for smartwatch localization using deep learning," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Nantes, France, Sep. 2018, pp. 1–8.

[25] B. Bhattarai, R. K. Yadav, H.-S. Gang, and J.-Y. Pyun, "Geomagnetic field based indoor landmark classification using deep learning," *IEEE Access*, vol. 7, pp. 33943–33956, 2019.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

[26] G. Wang, X. Wang, J. Nie, and L. Lin, "Magnetic-based indoor localization using smartphone via a fusion algorithm," *IEEE Sensors J.*, vol. 19, no. 15, pp. 6477–6485, Aug. 2019.

[27] S. Lee, S. Chae, and D. Han, "ILoA: Indoor localization using augmented vector of geomagnetic field," *IEEE Access*, vol. 8, pp. 184242–184255, 2020.

[28] B. Gozick, K. P. Subbu, R. Dantu, and T. Maeshiro, "Magnetic maps for indoor navigation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 12, pp. 3883–3891, Dec. 2011.

[29] N. Moayeri, M. O. Ergin, F. Lemic, V. Handziski, and A. Wolisz, "PerfLoc (Part 1): An extensive data repository for development of smartphone indoor localization apps," in *Proc. Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Sep. 2016, pp. 1–7.

[30] D. Hanley, A. B. Faustino, S. D. Zelman, D. A. Degenhardt, and T. Bretl, "MagPIE: A dataset for indoor positioning with magnetic anomalies," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2017, pp. 1–8.

[31] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, New York, NY, USA, 2015, pp. 178–189.

[32] J. Torres-Sospedra, D. Rambla, R. Montoliu, O. Belmonte, and J. Huerta, "UJIIndoorLoc-mag: A new database for magnetic field-based localization problems," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2015.

[33] *Information Technology—Real Time Locating Systems—Test and Evaluation of Localization and Tracking Systems*, Standard ISO/IEC 18305:2016(E), International Organization for Standardization, Geneva, CH, Standard, 2016.

[34] National Center for Health Statistics, "Advance data from vital and health statistics: Numbers 1-10," Nat. Center Health Statist., Hyattsville, MD, USA, Tech. Rep. 16(1), May 1989.

[35] D. Hanley, X. Zhang, A. S. D. de Oliveira, D. Steinberg, and T. Bretl, "Experimental evaluation of the planar assumption in magnetic positioning," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2018, pp. 1–8.

[36] R. Nicolai, "The premedieval origin of portolan charts: New geodetic evidence," *ISIS*, vol. 106, no. 3, pp. 517–543, Sep. 2015.

[37] E. R. Bachmann, X. Yun, and A. Brumfield, "Limitations of attitude estimnation algorithms for inertial/magnetic sensor modules," *IEEE Robot. Automat. Mag.*, vol. 14, no. 3, pp. 76–87, Sep. 2007.

[38] T. Harada, T. Mori, and T. Sato, "Development of a tiny orientation estimation device to operate under motion and magnetic disturbance," *Int. J. Robot. Res.*, vol. 26, no. 6, pp. 547–559, Jun. 2007.

[39] A. Das *et al.*, "Mapping, planning, and sample detection strategies for autonomous exploration," *J. Field Robot.*, vol. 31, no. 1, pp. 75–106, Jan. 2014.

[40] J. Jung, T. Oh, and H. Myung, "Magnetic field constraints and sequence-based matching for indoor pose graph SLAM," *Robot. Auto. Syst.*, vol. 70, pp. 92–105, Aug. 2015.

[41] B. Siebler, S. Sand, and U. D. Hanebeck, "Localization with magnetic field distortions and simultaneous magnetometer calibration," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3388–3397, Feb. 2020.

[42] B. Ferris, D. Hahnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *Robotics Science and Systems*. Philadelphia, PA, USA: MIT Press, Aug. 2006.

[43] R. Miyagusuku, A. Yamashita, and H. Asama, "Data information fusion from multiple access points for WiFi-based self-localization," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 269–276, Apr. 2019.

[44] I. Ashraf, M. Kang, S. Hur, and Y. Park, "MINLOC: Magnetic field patterns-based indoor localization using convolutional neural networks," *IEEE Access*, vol. 8, pp. 66213–66227, 2020.

[45] N. Akai and K. Ozaki, "3D magnetic field mapping in large-scale indoor environment using measurement robot and Gaussian processes," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2017, pp. 1–7.

[46] C. Jidling, N. Wahlström, A. Wills, and T. B. Schön, "Linearly constrained Gaussian processes," in *Proc. Adv. Neural Inf. Process. Syst.*, I. Guyon *et al.*, Eds. Red Hook, NY, USA: Curran Associates, Inc., 2017, pp. 1215–1224.

[47] B. Brzozowski and K. Kaźmierczak, "Magnetic field mapping as a support for UAV indoor navigation system," in *Proc. IEEE Int. Workshop Metrol. Aerosp.*, Jun. 2017, pp. 583–588.

[48] M. Kok and A. Solin, "Scalable magnetic field SLAM in 3D using Gaussian process maps," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 1353–1360.

[49] D. Caruso, A. Eudes, M. Sanfourche, D. Vissiere, and G. le Besnerais, "An inverse square root filter for robust indoor/outdoor magneto-visual-inertial odometry," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2017, pp. 1–8.

[50] D. Caruso, A. Eudes, M. Sanfourche, D. Vissiére, and G. le Besnerais, "Robust indoor/outdoor navigation through magneto-visual-inertial optimization-based estimation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 4402–4409.

[51] A. Leuzinger and A. Taylor, "Magneto-inductive technology overview," PNI Sensor Corp., Santa Rosa, CA, USA, PNI White Paper, Feb. 2010.

[52] N. Kim and T. Hawks, "Digital compass and magnetometer having a sensor coil wound on a high permeability and isotropic core," U.S. Patent 4 851 775 A, Jul. 25, 1989.

[53] T. J. Hawks, "Zero-offset magnetometer having coil and core sensor controlling period of an oscillator circuit," U.S. Patent 5 239 264, Aug. 24, 1993.

[54] X. Deng, Z. Zhang, A. Sintov, J. Huang, and T. Bretl, "Feature-constrained active visual SLAM for mobile robot navigation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 7233–7238.

[55] J. DeGol, T. Bretl, and D. Hoiem, "Improved structure from motion using fiducial marker matching," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 273–288.

[56] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3400–3407.

[57] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 573–580.

[58] E. Kayacan, Z. Zhang, and G. Chowdhary, "Embedded high precision control and corn stand counting algorithms for an ultra-compact 3D printed field robot," in *Proc. Robot., Sci. Syst.*, Pittsburgh, Russia, Jun. 2018.

[59] J. M. G. Merayo, P. Brauer, F. Primdahl, J. R. Petersen, and O. V. Nielsen, "Scalar calibration of vector magnetometers," *Meas. Sci. Technol.*, vol. 11, no. 2, p. 120, 2000.

[60] J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira, and B. Cardeira, "Geometric approach to strapdown magnetometer calibration in sensor frame," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 1293–1306, Apr. 2011.

[61] G. W. Legge, C. Downey, N. A. Giudice, and B. S. Tjan, "Indoor airport wayfinding for blind and visually impaired travelers," Federal Aviation Admin. Office Airport Saf. Standards, Washington, DC, USA, Tech. Rep. DOT/FAA/TC-TN16/54, Dec. 2016.

[62] K. Kanagu, K. Tsubouchi, and N. Nishio, "Colorful PDR: Colorizing PDR with shopping context in walking," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sapporo, Japan, Sep. 2017, pp. 1–8.

[63] C. Chun, T. M. Schmit, D. Dong, and H. M. Kaiser, "Economic evaluation of shelf-space management in grocery stores," *Agribusiness*, vol. 23, no. 4, pp. 583–597, Oct. 2007.

[64] R. Harle, "A survey of indoor inertial positioning systems for pedestrians," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1281–1293, 3rd Quart., 2013.

[65] D. Simon, *Optimal State Estimation*. Hoboken, NJ, USA: Wiley, 2006.

[66] T. D. Barfoot, *State Estimation for Robotics*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[67] G. M. Mendoza-Silva, J. Torres-Sospedra, and J. Huerta, "A more realistic error distance calculation for indoor positioning systems accuracy evaluation," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sep. 2017.

[68] R. Kumar, D. Castanón, E. Ermis, and V. Saligrama, "A new algorithm for outlier rejection in particle filters," in *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 2010, pp. 1–7.

[69] C. S. Maiz, J. Miguez, and P. M. Djuric, "Particle filtering in the presence of outliers," in *Proc. IEEE/SP 15th Workshop Stat. Signal Process.*, Aug. 2009, pp. 33–36.

[70] J.-S. Lee and W. Kyun Chung, "Robust particle filter localization by sampling from non-corrupted window with incomplete map," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 1133–1139.

[71] J. Back, *User Manual: RM3100 and RM2100 Geomagnetic Sensor*, document 1017252 R07, PNI Sensor Corporation, Circadian Way, Santa Rosa, CA, USA, Jun. 2016.

[72] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2006.

[73] GPy. (2012). *GPy: A Gaussian Process Framework in Python*. [Online]. Available: http://github.com/SheffieldML/GPy

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.

[74] J. Hensman, N. Fusi, and N. D. Lawrence, "Gaussian processes for big data," in *Proc. Conf. Uncertainty Artif. Intell.*, Bellevue, WA, USA, Jul. 2013, pp. 282–290.
[75] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. J. Julian, and M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 913–920.
[76] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 4th ed. Hoboken, NJ, USA: Wiley, 2007.

David Hanley (Member, IEEE) received the B.S. and M.S. degrees in aerospace engineering from the University of Illinois at Urbana–Champaign, Urbana, IL, USA, in 2013 and 2015, respectively, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering.

Augusto S. Dantas de Oliveira received the B.S. degree in aerospace engineering from the University of Illinois at Urbana–Champaign, Urbana, IL, USA, in 2019.

Since then, he works at Lockheed Martin, Sunnyvale, CA, USA, as a Guidance, Navigation, and Control Engineer supporting a variety of programs in the areas of system test, software engineering, and simulation analysis.

Xiangyuan Zhang (Member, IEEE) received the B.S. degree in computer engineering from the University of Illinois at Urbana–Champaign (UIUC), Urbana, IL, USA, in 2020, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering at UIUC.

His current research interests are control theory, game theory, reinforcement learning, and their intersections.

Dae Hyun (Alex) Kim received the B.S. degree in aerospace engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 2019.

He is currently working as a Backend Developer at a Fintech startup located in Chicago, IL.

Yusheng Wei received the B.S. degree in electrical engineering from the University of Illinois at Urbana–Champaign (UIUC), Urbana, IL, USA, in 2019.

He is currently a Robot Operating System (ROS) Engineer with Guangzhou Unipower Technology Company, Ltd., Guangzhou, China, focus on robot navigation and sensor integration.

Timothy Bretl (Senior Member, IEEE) received the B.S. degree in engineering and the B.A. degree in mathematics from Swarthmore College, Swarthmore, PA, USA, in 1999 and the M.S. and Ph.D. degrees both in aeronautics and astronautics from Stanford University, Stanford, CA, USA, in 2000 and 2005, respectively.

Subsequently, he was a Post-Doctoral Fellow with the Department of Computer Science, also at Stanford University. Since 2006, he has been with the University of Illinois at Urbana–Champaign, Urbana, IL, USA, where he is an Associate Professor of Aerospace Engineering and a Research Associate Professor with the Coordinated Science Laboratory.

Dr. Bretl received the National Science Foundation Faculty Early Career Development Award in 2010. He has also received numerous teaching awards at Illinois, including the AIAA Student Chapter Teacher of the Year Award in 2015, both the William L. Everett Award for Teaching Excellence and the Rose Award for Teaching Excellence in 2016, and both the College of Engineering Teaching Excellence Award and the Campus Award for Excellence in Undergraduate Teaching in 2018.

Authorized licensed use limited to: University of Illinois. Downloaded on June 04,2021 at 22:36:36 UTC from IEEE Xplore. Restrictions apply.