Multi-Modal Indoor Positioning of Mobile Devices

Joseph Menke
Department of EECS
U.C. Berkeley
Berkeley, USA
Email: joemenke@eecs.berkeley.edu

Avideh Zakhor
Department of EECS
U.C. Berkeley
Berkeley, USA
Email: avz@eecs.berkeley.edu

Abstract—In this paper we extend our previous results on WiFi and image localization to include magnetic sensing for multi-modal indoor localization. A two-step process is proposed that performs an initial localization estimate, followed by particle filter based tracking. Initial localization is performed using WiFi and image observations. For tracking we fuse information from WiFi, magnetic, and inertial sensors. We demonstrate the feasibility of this system using fingerprint maps that are collected with a single walkthrough of the building at normal walking pace. Further we reduce our database generation method from previous works to require only a smartphone and a foot mounted inertial measurement unit (IMU). Only a smartphone is needed for positioning after database generation. We present results for two locations: the Stoneridge Mall in Pleasanton, California, and the Doe Library at the UC Berkeley campus. We achieve an average location error of 2.6m across both locations.

I. INTRODUCTION

In recent years, indoor localization has received a great deal of attention [1]. It has a number of useful applications such as location-aware intelligent shopping assistants and indoor real-time navigation. However it is a technically challenging problem due to the lack of GPS signals indoors. Further, in practice it is important to use very little additional hardware and not require extensive modeling of building characteristics.

One approach to indoor positioning is to apply inertial dead reckoning using an inertial measurement unit commonly found in today's mobile devices [2]. A commonly adopted method is to detect steps and then estimate the corresponding step lengths [3], [4]. A more accurate method is to use a foot-mounted sensor to track movement [5]. This utilizes the zero-velocity state of the foot when walking to estimate biases of the sensor and provides accurate positioning but requires additional hardware.

Another common approach is to use the WiFi infrastructure that is already prevalent inside most modern buildings. A popular method for utilizing this infrastructure is to construct a database of WiFi Received Signal Strength Indicator (RSSI) fingerprints for the building [7], [8]. A major advantage of this method is its prevalence as hardware infrastructure and the ubiquity of WiFi scanning capability on mobile phones and consumer electronic devices. A disadvantage is that the location dependency of RSSI is subject to interference and signals can be very similar in wide-open spaces.

Recently an image-based indoor localization scheme [9], [10] has been proposed for mobile devices with cameras using a locally georeferenced database of images. The Scale Invariant Feature Transform (SIFT) [11] allows accurate matching of images from a client side mobile device with those contained in the database. Its performance is degraded when the query image has few distinguishing features, or when the pictures are of low quality due to out-of-focus and/or motion blur.

Lastly there have been recent efforts to perform localization using distortions in the earths magnetic field caused by metal structures in buildings. If these distortions are mapped, it is possible to use them for localization [12], [13]. The advantages to using magnetic data are the prevalence of these sensors in mobile devices and that no new hardware infrastructure is required. One disadvantage though is that it takes about 10m of walking to generate a reasonable estimate of position, which is undesirable for many applications.

In this paper we extend our previous results on WiFi and image localization to include magnetic sensing for indoor localization. The outline of this paper is as follows: In Section 2 we present a method of database generation combining foot mounted Inertial Measurement Unit (IMU)-based pedestrian dead reckoning (PDR) and a mobile device. In Section 3 we propose a two-step localization initialization and realtime tracking method that combines information from magnetic, WiFi, accelerometer, gyroscope, and image-based sensors. In Section 4 we show results for two locations; the Stoneridge Mall in Pleasanton, California, and the Doe Library on the UC Berkeley campus.

II. DATABASE GENERATION

To generate our fingerprint database a smartphone is carried throughout the building in a single walkthrough at normal walking pace. A single foot mounted IMU that is capable of tracking movement over long distances is used to determine the location of the smartphone during data collection. The phone is held in front of the user with a constant orientation so a known offset with respect to the direction of walk can be determined. Positions recorded from this system are manually aligned with a floor plan to provide fingerprint locations in the common coordinate frame.

We use the technique described in [16] and [10] to collect wifi and image databases respectively. As the fingerprinting smartphone is carried through the building we record scans that consist of access point MAC addresses and the observed

signal strength for each access point. A scan contains measurements for all access points viewable by the device during the duration of the scan. Each scan is associated with a single location in the map. This results in a fingerprint which is stored in a SQL database on a local server. The image database is made up of images, camera positions, and sparse depth maps [10]. To collect images during the walkthrough the phone is held vertically with the camera pointing perpendicular to the direction of walk. Images are collected at a frequency of roughly 1Hz. A sparse depth map is then computed for each image using the method proposed in [10].

A magnetic database consists of observed magnetic vectors that have been rotated back to a global coordinate frame. We average sensor readings around positions reported by the foot mounted IMU in order to obtain the observation at that position. Then we perform tilt-compensation using the onboard accelerometer. The tilt-compensated magnetic vector is determined by:

$$c_x = m_x cos(a_y) + m_y sin(a_x) sin(a_y) - m_z cos(a_x) sin(a_y) \quad (1)$$
$$c_y = m_y cos(a_x) - m_z sin(a_x) \quad (2)$$
$$c_z = m_x sin(a_y) + m_z cos(a_x) cos(a_y) - m_y cos(a_y) sin(a_x) \quad (3)$$

where $\vec{m} = [m_x, m_y, m_z]^T$ is the original magnetic vector, a_x and a_y are the acceleration values of the x and y axis, and $\vec{c} = [c_x, c_y, c_z]^T$ is the tilt compensated magnetic vector. This allows us to remove effects of minor changes in pitch and roll that can occur while the user walks through the building.

To obtain the magnetic vector in the database coordinate frame we also need to determine the yaw of the phone. While the phone reports an orientation from its IMU, this orientation is dependent on the magnetic vector at that location, making it an inappropriate choice for determining the pose with respect to the database coordinate frame. Rather we opt to infer the yaw of the phone from successive positions reported by the foot mounted IMU.

III. Global Localization and Tracking

We fuse two sensing modalities, WiFi and Image, to arrive at the global location initialization. For the image modality the user takes a picture of the area around them. This image is then sent to our server for pose computation using the method presented in [10]. This finds a matching image in the database and computes the camera pose of the query image. For WiFi, the initial estimate is determined using the clustering method proposed in [16]. The centroid of the cluster is returned representing an estimate of the position of the mobile device.

We use a particle filter based tracking method whereby the particles state vector consists of the x and y position and yaw orientation of the particle [15]. The initial locations of the particles are samples from a 2D Gaussian distribution with a mean equal to that of the location estimate provided by either WiFi or image based initialization. If both WiFi and image initialization report reasonable confidence values then half the particles are sampled from each distribution. Initial yaw estimates are sampled uniformly in the range [0,360). The location reported by the filter at each interval is the weighted average position of all particles.

For the propagation of the particle filter we perform step detection on the mobile device and propagate the particles at each step. Algorithms presented in [17] are used to detect steps and estimate their lengths from accelerometer readings. A change in yaw is determined from gyroscope measurements. For each particle, we add a random noise to both the rotation and translation of each step. If a step causes a particle to cross a wall in the floorplan then it is assumed that the particle cannot have represented a true location and it is eliminated.

WiFi scans are collected continuously as the user walks. At the completion of each scan the observations are sent to the server where the normalized Redpin score for each fingerprint is computed. When a response is received by the server particles are weighted by the score of the fingerprint closest their position at the time at which the scan was recorded.

The onboard digital compass of the mobile device is used to record the magnetic vector. Though we assume the phone is held at a fairly constant orientation, we know that it would be impossible to keep the phone perfectly stationary and so we perform tilt compensation on the magnetic reading using the same method as in the database generation. For each particle we then find the closest fingerprint in the database using a quad-tree search. The magnitude of the observation is compared to the magnitude of the closest fingerprint. The probability is assumed to be Gaussian centered at the magnitude of the fingerprint. This represents the probability that a particle is at the correct position.

To determine the probability that a particle is at the correct orientation we project both the observed magnetic vector and the database magnetic vector to the X-Y plane. Since this value has already been tilt compensated, this projection is merely the X and Y components of the tilt-compensated vector. From these values we can estimate orientation of the particle from the measurement:

$$\theta = tan^{-1}((\vec{v_o} \times \vec{v_d})/(\vec{v_o} \cdot \vec{v_d})) \quad (4)$$

where $\vec{v_o}$ and $\vec{v_d}$ are the projections of the observed and database magnetic vectors respectively and θ is the estimated orientation of the particle. As seen in Fig. 1, we compute the difference between the estimated orientation, determined from the yaw angle between the database vector and the observed vector, to the orientation of the particle. If this difference is large, represented by particle "A", the particle is assigned a low probability. A small difference, represented by particle "B", results in a high probability. The probability is assumed to be Gaussian centered at the expected orientation. The particle is weighted by the product of the location and orientation probabilities.

IV. Experimental Results

We test our system in two locations. The first is Stoneridge Mall in Pleasanton, California. The second is the Doe Library
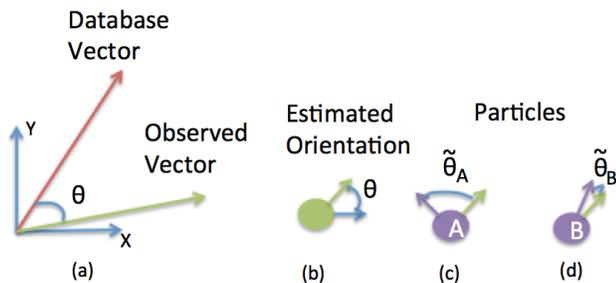
Fig. 1. Method for weighting particles based on Magnetic Observation; the green arrow is the devices estimated orientation and the purple arrow is the particles orientation. (a) The database vector and the observed vector are used in Equation 1 to estimate orientation. (b) Estimated orientation θ of the device. (c), (d) Comparison of two hypothetical particles. Particle "B" would be weighted with a high probability since the difference from the estimate, denotes $\tilde{\theta}_B$, is small. Particle "A" would receive a low probability since the difference from the estimate is quite large.

on campus at UC Berkeley. For both test cases fingerprint data is collected using a smartphone and a foot mounted PDR device. Data collection is performed by one person in a single walkthrough of the building at normal walking pace.
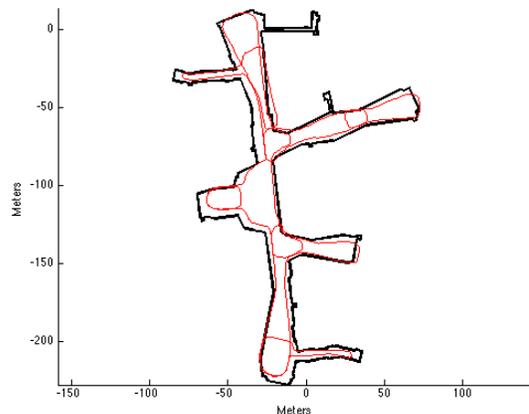
To characterize performance the evaluator wears the same foot mounted PDR device in order to compare the location estimate to a "ground truth". Errors are computed based on the position error of our system every time a new location is reported by the PDR device, which is about once per second. While we denote the location from the PDR device as a "ground truth", the locations reported by this device are error prone requiring the the "ground truth" path to occasionally go outside our floorplan, increasing, possibly significantly, the estimate of our error.
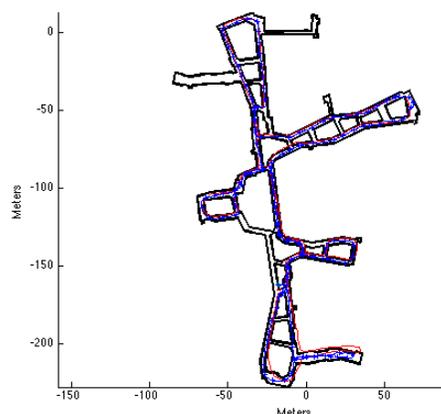
TABLE I
EXPERIMENTAL RESULTS FOR THE STONERIDGE MALL.

	Err. Mean(m)	Err. Std.(m)	90% Err.(m)	Length(m)
D1	2.15	1.21	3.7	830
D2	2.19	1.45	3.8	827
D3	2.92	1.41	5.0	835
D4	2.70	1.21	4.3	838
R1	2.64	2.69	4.0	921
R2	3.03	2.98	5.4	909
R3	2.61	1.51	4.6	892
R4	2.63	1.39	4.6	1022
R5	2.74	2.13	5.8	1006
R6	2.04	1.42	3.9	711
Avg.	2.56	1.74	4.51	879

A. Stoneridge Mall

For the Stoneridge Mall data for the database is collected using a Samsung Galaxy S4 smartphone. The collection path is 1417m as shown in Fig. 2(a). For testing a Google Nexus 5 phone is used to walk ten paths, each of which spanned a majority of the mall. Table I shows results of the individual trials. Of these ten paths, four are identical with the same start location and path walked. These paths are labeled "D". The remaining six paths each have different start locations and trajectories. These are labeled "R". Fig. 2(b) shows an example

(a)

(b)

Fig. 2. Example test paths in Stoneridge Mall. Red is the "ground truth" as measured by the foot mounted IMU. Blue is the path estimated by our system. 'x's indicate locations where a WiFi update is received. (a) Database collection path, (b) Path R5 from Nexus 5.

of a "R" path. Average length of paths is about 880m with an average position error of about 2.6m. There was no significant difference in error between "D" and "R" paths.

For this experiment the database was collected over a month prior to final testing. This speaks to the stability of our maps, showing that a single walkthrough is capable of generating maps that are valid long after the collection was performed. Tests were performed on a Saturday, when the mall was very busy, over the course of an entire day.

There are two main contributors of error for this experiment. The first is that one of the two initialization methods may return a location that is sometimes erroneous. This results in a position being reported between the two initialization locations until the incorrect particles die off. The other significant contributor to error in this experiment is that our "ground truth" method is not quite accurate for such a large area. Several areas report significant error due to the "ground truth" being reported in areas outside our floorplan.
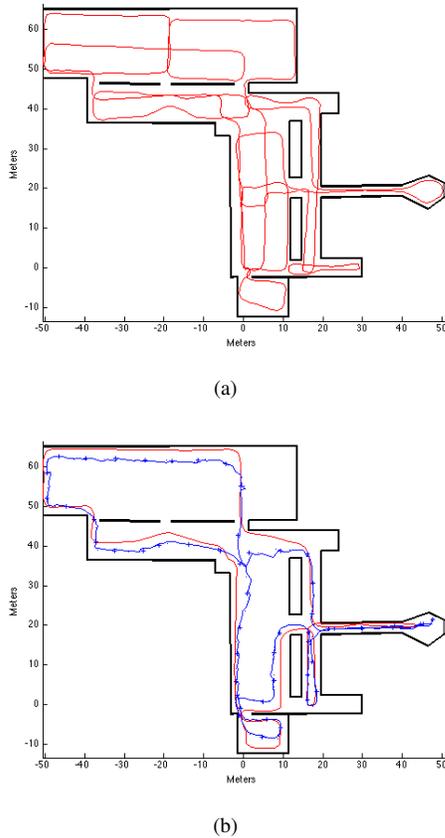
(a)

(b)

Fig. 3. Example test paths in Doe Library. Red is the "ground truth" as measured by the foot mounted IMU. Blue is the path estimated by our system. 'x's indicate locations where a WiFi update is received. (a) Database collection path, (b) Path D2 from Nexus 5

TABLE II
EXPERIMENTAL RESULTS FOR THE DOE LIBRARY.

	Err. Mean(m)	Err. Std.(m)	90% Err.(m)	Length(m)
D1 N5	2.09	1.11	3.8	383
D2 N5	2.17	1.45	3.8	381
R1 N5	2.42	1.68	5.1	418
R2 N5	3.11	2.42	6.6	361
R3 N5	2.11	1.57	3.6	372
D1 S4	2.45	1.33	4.2	382
D2 S4	3.50	2.55	7.7	380
R1 S4	2.62	1.89	4.3	373
R2 S4	2.81	2.09	4.8	408
R3 S4	3.18	3.25	5.7	352
Avg.	2.65	1.93	4.96	381

B. Doe Library

For the Doe library data for the database is collected using a Google Nexus 7 tablet. The collection path is 880m as shown in Fig. 3(a). For testing a Google Nexus 5 phone and a Samsung Galaxy S4 phone were each used to walk five paths for a total of ten paths. Table II shows the results of the ten trials. Trials taken with the Google Nexus 5 and Samsung Galaxy S4 are labeled "N5" and "S4" respectively. All paths labeled "D" were taken along the same route as shown in Fig. 3(b). Paths labeled "R" are all different with different starting points. Average length of paths is about 380m with an average position error of about 2.7m. There was no significant difference in error between "D" and "R" paths.

While our errors are in line with those obtained at the Stoneridge Mall, drift of particles towards the center of the room due to the low magnetic and WiFi signal variation is likely a larger contributor to error in this location than errors in the "ground truth".

REFERENCES

[1] H. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems", *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE*, Volume:37 , Issue: 6, 2007

[2] R. Harle, "A Survey of Indoor Inertial Positioning Systems for Pedestrians", *Communications Surveys & Tutorials, IEEE* Volume: 15, Issue 3, 2013

[3] M. Holcik, "Indoor Navigation for Android," *Masters Thesis, Faculty of Informatics, Masaryk University*, 2012.

[4] Y. Jin, H.Toh, W. Soh, and W. Wong, "A robust dead-reckoning pedestrian tracking system with low cost sensors," *IEEE international Conference on Pervasice Computing and Communications (PerCom), Seattle, WA, USA*, 2011

[5] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors", *Computer Graphics and Applications, IEEE* , Volume:25 , Issue: 6, pp 38-46, 2005

[6] P. Bolliger, "Redpin - Adaptive zero configuration indoor localization through user collaboration",*Workshop on Mobile Entity Localization and Tracking in GPS-less Environment Computing and Communications Systems (MELT), San Francisco* 2008

[7] B. Ferris, D. Fox, N. Lawrence, "WiFi-SLAM Using Gaussian Process Latent Variable Models", *IJCAI* 2007

[8] P. Bolliger, K. Partridge, M. Chu, and M. Langheinrich, "Improving location fingerprinting through motion detection and asynchronous interval labeling",*Lecture Notes in Computer Science* Volume 5561, pp 37-51 2009

[9] A. Hallquist and A. Zakhor, "Single view pose estimation of mobile devices in urban environments," *IEEE Workshop on the Applications of Computer Vision (WACV), Clearwater, FL, USA*, 2013

[10] J. Z. Liang, N. Corse, E. Turner, and A. Zakhor, "Reduced-complexity data acquisition system for image-based localization in indoor environments," *International Conference on Indoor Positioning and Indoor Navigation (IPIN '13), Montbeliard-Belfort, France*, 2013

[11] D. G. Lowe, "Distinctive image features from scale invariant key points", *International Journal of Computer Vision*, pp 91-110 2004

[12] J. Haverinen, "Global indoor self-localization based on the ambient magnetic field",*Journal of Robotics and Autonomous Systems*, Volume 57 Issue 10, pp 1028-1035 2009

[13] J. Chung, M. Donahoe, C. Schmandt, I. Kim, P. Razavai and M. Wiseman, "Indoor location sensing using geo-magnetism", *9th international conference on Mobile systems, applications, and services* 2011

[14] J. Zhang, A. Hallquist, E. Liang, and A. Zakhor, "Location-Based Image Retrieval for Urban Environments," *in International Conference on Image Processing*, 2011.

[15] P. Levchev, M. Krishnan, C. Yu, J. Menke, and A. Zakhor, "Simultaneous Fingerprinting and Mapping for Multimodal Image and WiFi Indoor Positioning." IPIN 2014, Busan, Korea, October 2014.

[16] P. Levchev, C. Yu, M. Krishnan, A. Zakhor, "Indoor WiFi Localization with a Dense Fingerprint Model" UC Berkeley, Tech. Rep. Apr. 2014 [Online]. Available: http://www-video.eecs.berkeley.edu/papers/plamen/globe14-localization-0401.pdf

[17] A. Serra, T. Dessi, D. Carboni, V. Popescu, and L. At- zori, "inertial navigation systems for user-centric indoor applications,"*Networked and Electronic Media (NEM '10) Summit, Barcelona, Spain*, 2010.