Journal of Engineering Sciences
Vol 11, Issue 4 ,April/2020
ISSN NO:0377-9254

Indoor Navigation system based on vision for smartphones

P. Phanindra kumar Reddy, J. Mounika, I. Naga Hanoosha, D. Mrunalini, B. Lavanya

Computer Science and Engineering

Annamacharya Institute of Technology and Sciences, Rajampet,India

ABSTRACT:

Indoor navigation services focused on smartphones are so desperately needed in indoor environments. Nevertheless, their adoption was relatively slow, due to the lack of fine-grained and up-to-date indoor maps, or the potentially high implementation and maintenance costs of indoor localization solutions based on infrastructure. This work proposes ViNav, a scalable and cost-effective system that uses visual and inertial sensor data collected from smart phones to implement indoor mapping, localization and navigation. ViNav uses structure-from-motion (SfM) techniques to reconstruct indoor 3D models from crowd-based images, locate points of interest (POIs) in 3D models, and compile 3D models. ViNav introduces image-based localization that defines the locations and directions of the users, and leverages this function to calibrate dead-reckoning user trajectories and sensor fingerprints collected along the trajectories.The optimized information is used to create indoor maps more descriptive and precise, and to reduce the response delay of requests for localisation. The system functions properly according to our experimental results in a university building and a supermarket, and our indoor localization achieves competitive performance compared to traditional approaches: in a supermarket. ViNav locates users in 2 seconds, with less than 1 meter of distance error and less than 6 degrees of direction error in the face.

I. INTRODUCTION:

Indoor mobile navigation systems are essential for complex indoor environments such as airports, malls and museums. Unfortunately, the in-door navigation systems adoption rate is still very low, although initial efforts were made several decades ago to deploy them. The slow progress has been attributed to several factors. First, indoor navigation requires indoor maps that are fine-grained and up-to-date to calculate navigation routes and search for points of interest (POI). Many solutions1 ask users to show images of floor plans as their indoor maps in public venues. Most of them, however, lack data, or were not kept up to date. So they can never be used directly to navigate indoors. Secondly, even in the case of specific indoor maps, most systems rely on pre-scanned radio maps or pre-installed hardware [14],[19] for location, which are expensive to install and difficult to maintain. So, we saw it worth investigating whether we could build an alternative method of indoor navigation that would not require pre-created indoor maps, pre-scanned radio maps or pre-installed hardware.

Today, smartphones are fitted with high-resolution cameras, and mobile users are willing to share some of their images in public, e.g. through photo-sharing websites such as Flickr or Instagram. In addition, researchers have demonstrated that crowdsourced images can be used to create 3D models of indoor and outdoor environments using structure-from-motion (SfM)4 techniques. For SfM-based 3D models, the development of maps and further navigation meshes has good potential. In addition, image-based localization services can be made available based on matching features. Inspired by this, we are proposing to create a system, ViNav, that uses crowdsourced visual data and SfM— based 3D models to solve indoor mapping, localization and navigation problems as a whole.

Our system is based on sensor-enriched 3D models which are bootstrapped and upgraded from smartphones using crowdsourced visual and sensor data collected. By using SfM techniques, ViNav first takes crowdsourced images and videos as an input to create 3D models (in the form of 3D point clouds) of an indoor area of interest. It then creates navigation meshes for path planning based on the information obtained from 3D models about obstacles. Because the crowdsourced photos are taken at arbitrary locations, the generated 3D model may not cover every aspect of a space,and may suffer from uneven distribution of density points. That could result in incomplete meshes of navigation. To overcome this issue, we suggest using motion sensors to monitor

user trajectories and incorporate pedestrian paths derived from crowdsourced user trajectories into the navigation meshes. We also allow multi-storey building navigation by using barometric pressure sensor readings when users use our system to detect connecting paths and link 3D.ViNav also extracts POI data from the collected visual data and maps it to 3D space, making it easy for users to locate POIs. For 3D models based on SfM, the system provides image-based localization that uses photos taken in situ to define the user's location and direction of face. The problem with this approach is that the response time increases as the sizes of the 3D models increase, because position is calculated by comparing 2D features in a query picture against 3D points found in the 3D models. ViNav uses model partitioning based on the density of 3D points to enable fast localization, and selects partitions based on Wi-Fi fingerprinting to match features. Experimental results show decreased delays in the response to localization for this partitioning scheme even for large 3D models. The system can provide a navigation route and direct mobile users to their destinations by using the image-based localisation, the navigation mesh, and the defined POI location.

We present the concept, implementation and detailed evaluation of ViNav in this paper. The paper's findings are summarised as follows:

• To allow ViNav to tackle several technical challenges: (1) From crowdsourced user trajectories ,we combine recognized pedestrian paths to complement crowdsourced indoor 3D models; (2) we fuse data of barometric pressure with user trajectories which helps in detecting connecting paths between floors; (3) extracting POI information from crowdsourced visual data and translating the extracted POI into 3D space; (4) accelerating the localization process by density-based model partitioning and fingerprint-based partition selection.

• To determine its feasibility and performance, we introduce a prototype of ViNav. Experimental results show that ViNav performs well in multi-storey buildings and shows good efficiency. ViNav locates a user within 2 seconds, with a location error of less than 1 meter and a facing direction error of less than 6 degrees, according to our experiments conducted in two distinct indoor environments.

The ViNav indoor navigation system meets the requirements set out above. The system is based upon the architecture of a client / server. The application serves two purposes: (1) collecting visual data and other sensing data, and uploading the collected data to the server to create sensor-enriched 3D indoor environment models; (2) providing POI search interfaces; (3) identifying positions and presenting users with simple navigation pathways.

II.INDENTATIONS AND EQUATIONS

EVALUATON

1.Methodology

ViNav can be used for locating users, searching for places of interests, calculating navigation routes, and providing AR navigation guidance. Performance of *ViNav* depends on the performance of the localization, the accuracy of recognized PoIs, and the accuracy of the navigation meshes used for path planning. Because the navigation meshes are generated from the sensor-enriched 3D models (cf. Section 3), the accuracy of the navigation meshes is determined by the quality of the 3D models in use. Similarly, the accuracy of the navigation depends on the accuracy of the estimates of user's position and facing direction. Therefore, for performance evaluation of *ViNav*, we measure the following metrics: (1) the accuracy of the indoor layout extracted from the 3D models, (2) the accuracy of detecting stairs and elevators in multi-floor buildings, (3) the performance of POI detection, and (4) the performance of the 3D-model- based indoor localization.

2. Accuracy of Navigation Mesh

We initially assess the precision of the indoor format separated from 3D models. We gathered information in two distinctive indoor conditions. The first is a place of business, where Department of Computer Science at Aalto University finds (we name it as CS fabricating from there on). It comprises of 3 stories. During weekdays, in excess of 500 understudies and staffs visit the CS expanding on regular routine. The ground floor is around 1,100 m2, covering a long passageway, a library, furthermore, a cafeteria. The indoor scenes of the cafeteria change frequently because of advancements and designs. A development work of another nourishment store was begun during the field study. The subsequent structure is a major store

in Helsinki. Its size is around 1,600 m2. There are in excess of 30,000 items selling in the grocery store.

2.1 CS Building

ViNav can be utilized for finding clients, scanning for spots of interests, figuring route courses, and giving AR route direction. Execution of ViNav relies upon the exhibition of the confinement, the precision of perceived PoIs, and the exactness of the route networks utilized for way arranging. Since the route networks are produced from the sensor-improved 3D models (cf. Area 3), the exactness of the route networks is controlled by the nature of the 3D models being used. Also, the precision of the route relies upon the exactness of the evaluations of client's position and confronting bearing. In this manner, for performance assessment of ViNav, we measure the accompanying measurements: (1) the precision of the indoor format extricated from the 3D models, (2) the exactness of recognizing stairs and lifts in multi-floor structures, (3) the presentation of POI identification, and (4) the exhibition of the 3D-model-based indoor limitation.

To assess the precision of the remade plot, we determined the Euclidean separations of dividers encompassing the CS building dependent on the 3D models, and estimated the ground truth utilizing a laser rangefinder. From that point onward, we determined the distance blunder as characterized in [28]. In particular, the separation mistake for an accurately distinguished section is characterized as the normal separation between the recognized fragment and its related ground truth portion. The general separation blunder, meant by ED, is determined beneath.

$$E_D = \sum_{i=1} [length(s_i) E_d(s_i)] / \sum_{i=1} length(s_i) \quad (1)$$

Where N is a number of correctly detected segment s and where Ed(si) is the distance error for a correct ly detected segment.

ViNav compiles a navigation mesh based on the revised 3D model with 215 obstacles. The Navigation Mesh is shown in Fig. 8b. 8b. According to Compare, we calculate the mean position error of obstacles. (2). (2). Denote the barrier coordinates in the mesh by (xi, yi) and its actual coordinates by (Xi, Yi), the average barrier position error, denoted by obstacles, can be calculated as

$$E_{obstacles} = \sum_{i=1}^{N} (x_i - X_i)^2 + (y_i - Y_i)^2 / N \quad (2)$$

Where N represents the number of obstacles. As a r esult, the average position error is 0.78 metres, sug gesting the navigation mesh's reasonably high accur acy.

III. Figures

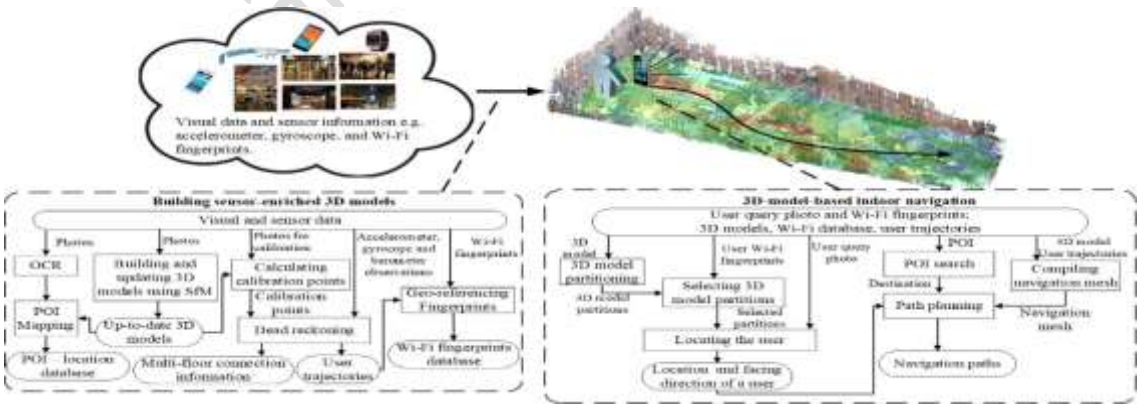
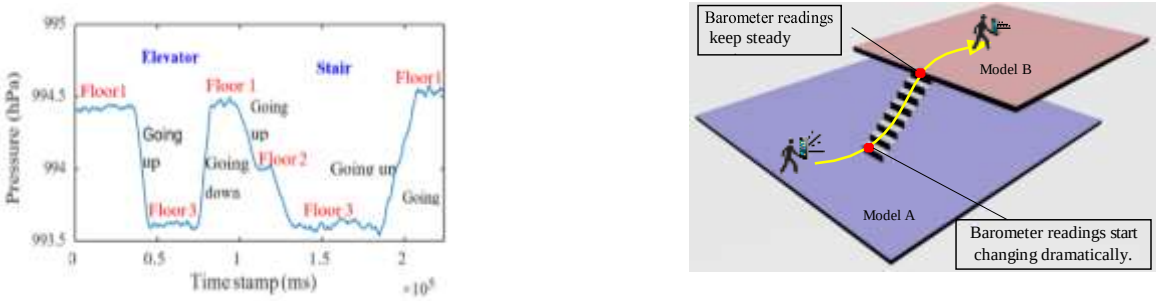


Fig. 1: Overview of a mobile crowdsensing-based indoor navigation system for smartphone.

Journal of Engineering Sciences
Vol 11, Issue 4 ,April/2020
ISSN NO:0377-9254



(a) (b)

Fig. 2: Detecting transitions between floors based on barometer readings. (a) Pressure variations when changing floors. (b) Marking a connecting path (red dots) between floors.

Fig. 3: An example of locating a recognized word. The dashed rectangle indicates the area where a word is recognized. The dots represent the feature points extracted from the image. The circle shows the feature point which is the closest one to the center of the recognized word.

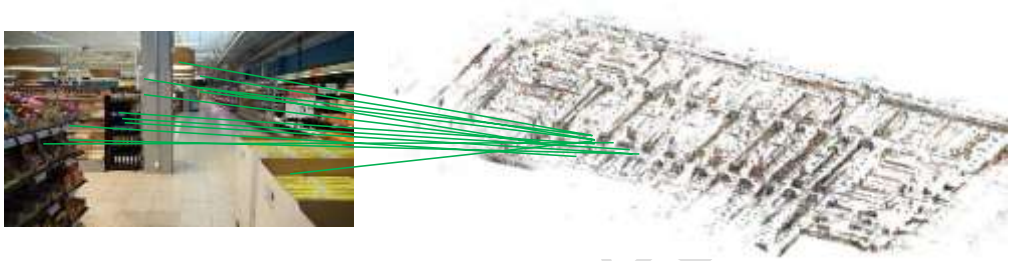


Fig. 4: Each feature point extracted from the input image has a corresponding 3D position in the SfM point cloud.

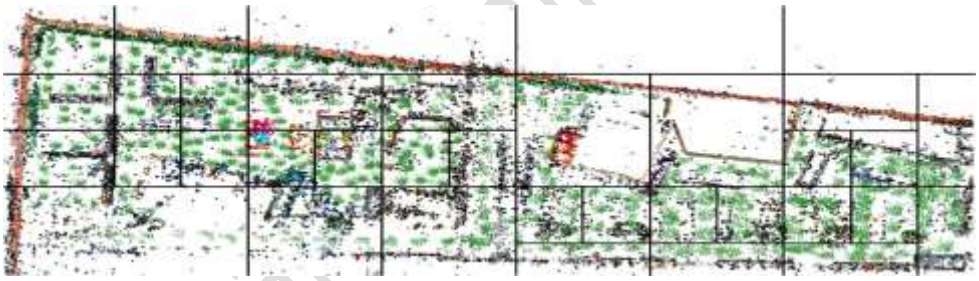


Fig. 5: An example of density-based point cloud partitioning. Each partition includes points corresponding to features extracted from no more than 100 photos. Both the width and length of each partition are larger than 5 meters. The camera positions are marked as green dots in the figure.

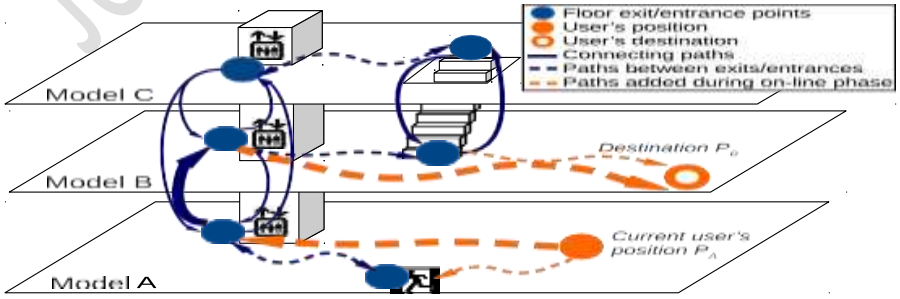
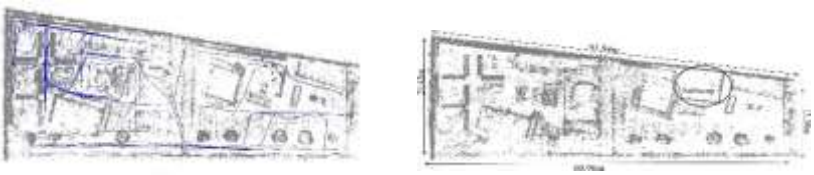


Fig. 6: Multi floor navigation graph for a 3 storey building. Blue nodes (entry and exit points) and edges (connecting and navigation paths between them) represent graph *G*. Orange parts are automatically added before computing a path between user's position and destination. The thicker edges represent the shortest path.



(a) (b) (c)

Fig. 7: Comparison of the generated 3D models with the actual layout of the CS building. (a) Official floor plan. The layout of shelves in the library has changed since the plan was released. (b) A 2D view of the initial 3D model. The blue lines marked on the figure indicate the walking traces of the volunteers. (c) A 2D view of the updated 3D model.

(a) (b)

Fig. 8: Navigation meshes generated from 3D models. (a) Navigation mesh of the library area. The grey solid polygons represent the one compiled from the initial 3D model only, while the black hollow polygons correspond to the one utilizing also the user trajectories. (b) The navigation mesh of the CS building updated in March, 2015.

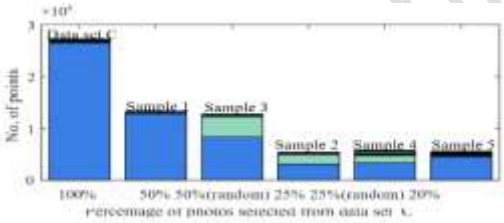


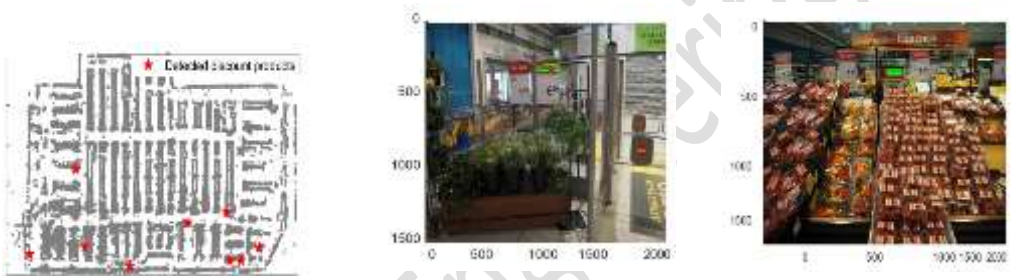
Fig. 9: Comparison of the number of 3D points included in the 3D models built from different samples. The different color in each bar represents the number of points included in each fragment of the 3D model in question.

(a) (b) (c)

Fig. 10: 3D modelling result of the supermarket. (a) Floor plan of the supermarket. (b) A 2D view of the reconstructed point cloud. (c) A comparison between the floor plan and the point cloud.

Staircase name	No. of measurements	Position error (m)
Floor 1		
Staircase to 2nd floor	5	1.3622
Staircase to 2nd floor Elevator	4	2.5137
	15	4.2298
Floor 2		
Staircase to 1st floor	6	1.8135
Staircase to 3rd floor	13	2.6886
Staircase to 1st floor	18	1.3317
Staircase to 3rd floor	8	0.8875
Elevator	3	0.5212
Floor 3		
Staircase to 2nd floor	12	2.1184
Staircase to 2nd floor	6	4.6336
Elevator	8	1.3539

TABLE 1: Discovered connecting paths (stairs and elevators) in a 3 storey building. The second column shows how many detected positions were used in calculating a connecting path endpoint. The last one shows an Euclidean distance error between an estimated endpoint and a ground truth.



(a) (b) (c)

Fig. 11: Examples of POI mapping. The red stars in (a) refer to the positions of the sale signs in the supermarket. (b) and (c) show example images from which the sale signs are recognized by text and marked with green rectangles.

(a) (b)

Fig. 12: Measurements points in two environments. (a) 185 measurement points in CS building. They were distributed in the cafe (black circle), corridor (red triangle) and library (blue square). (b) 71 measurements in the supermarket.

Fig. 13: Evaluation toolkit. The left-hand figure shows a self- built toolkit that consists of a laser rangefinder, a protractor, an Android phone running *ViNav* mobile app, and a tripod as a stand. The right-hand figure demonstrates how trilateration is applied to determine the location of the toolkit based on measurements of the distance to each reference point.

Area	No.of Measurement Points	No.of Photos	Hit rate		
			Image	**Wi-Fi**	***ViNav***
Café	78	936	91%	100%	91%
Corridor	37	444	100%	100%	100%
Library	70	840	98.6%	100%	97.1%
Total	185	2,220	95.7%	100%	95.1%

TABLE 2: Comparison of hit rate among image, Wi-Fi and *ViNav* in the CS building.

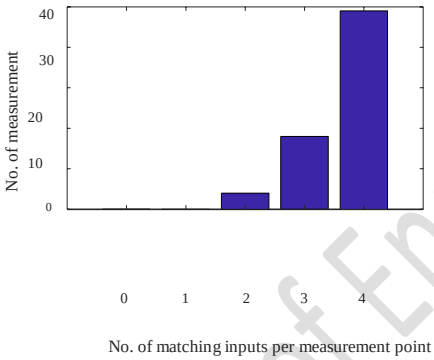


Fig. 14: 69% of the measurement points in the supermarket have 4 matching inputs, while 25.4% and 5.6% of them have 3 and 2 matching inputs respectively.

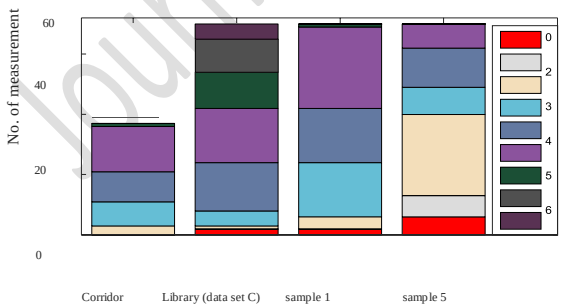
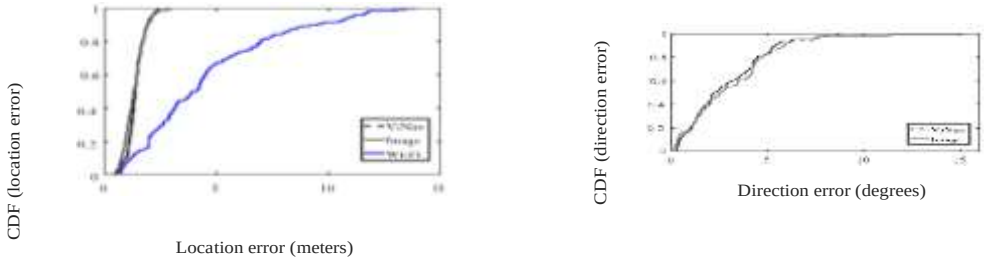


Fig. 15: The numbers of measurement points where the indicated amount of matching inputs are identified at each point in the CS building. Sample 1 and 5 are subsets of data set C in the library. Each color represents the number of matching inputs at a certain measurement point. For example, there are 15 (out of 37) measurement points in the corridor. Each point has 6 matching inputs.



(a) **(b)**

Fig. 16: Comparison of (a) location error and (b) direction error between *ViNav*, Image, and Wi-Fi in the CS building. Note that Wi-Fi does not return user's facing direction.

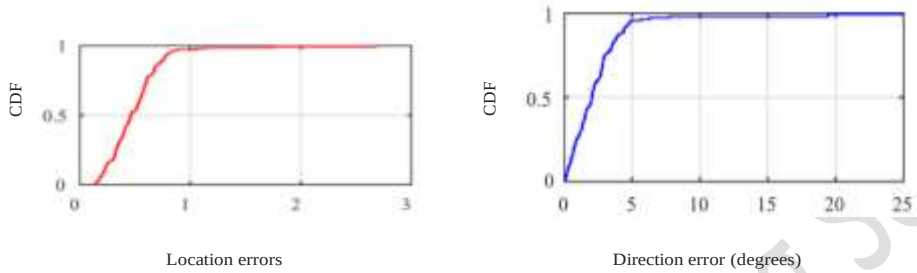


Fig. 17: (a) Location error and (b) direction error of *ViNav* in the super market.

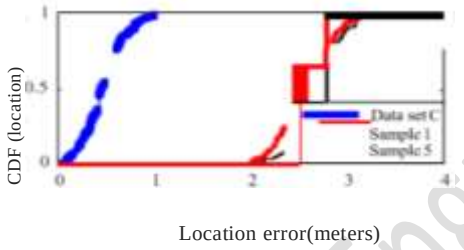
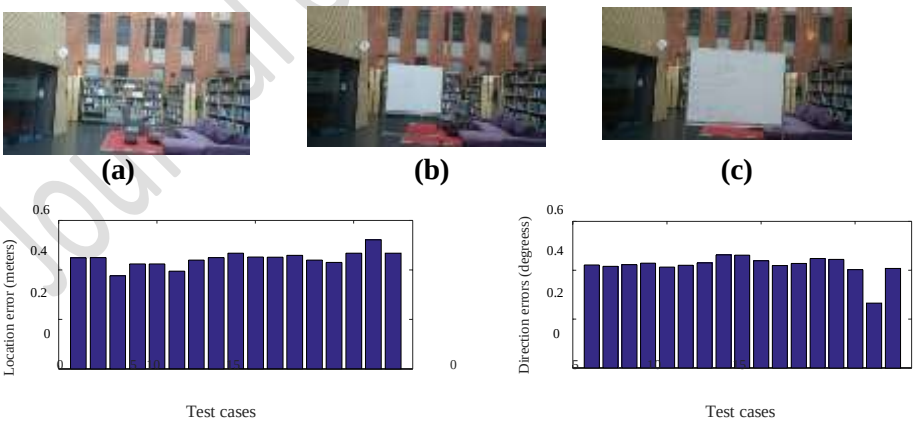


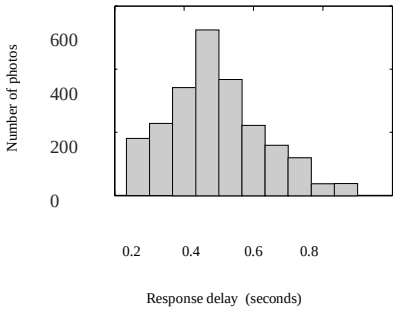
Fig. 18: Comparison of location error between scenarios where the 3D models were built from data set C, Sample 1, and Sample 5, respectively.



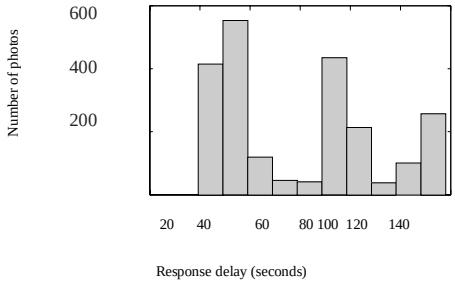
(a) **(b)** **(c)**

(d) **(e)**

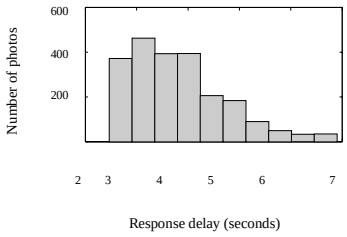
Fig. 19: Location and facing direction errors in the cases where a whiteboard was brought into the scene and then moved towards the camera.

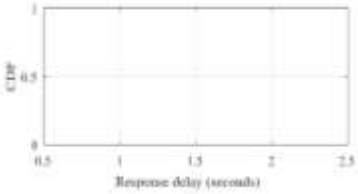


(a) Wi-Fi Fingerprinting



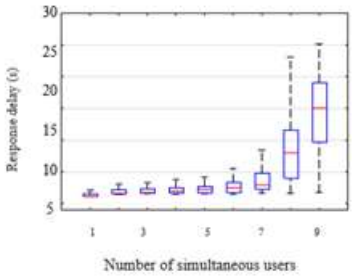
(b) Image-based localization



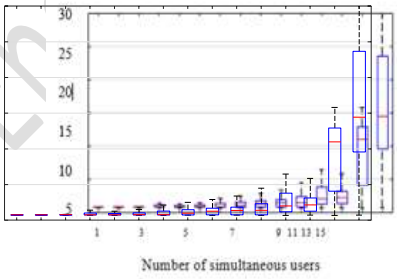


(d) ViNav (Supermarket)

Fig. 20: Comparison of response delay between indoor localization algorithms in the CS building (a)(b)(c) and supermarket (d).



(a) *ViNav* on a public cloud



(b) *ViNav* on an edge server

Fig. 21: *ViNav* response delays for different numbers of simultaneous users. The marking inside each box shows the median, bottom and top edges of the box indicate the 25th and 75th percentiles, the whiskers extend to the most extreme data points.

IV. Conclusion

In this paper we presented ViNav, a low-cost and entire-system solution for indoor navigation. ViNav partly builds on several existing techniques such as SfM and fingerprinting. Nonetheless, ViNav brings new functions and tackles many technological challenges to enable indoor navigation based on mobile crowdsensing. It uses crowd-based visual information to create 3D models of fascinating indoor spaces, and detects pedestrian paths from crowd-sourced user trajectories. By taking advantage of the high hit rate and low response time of Wi-Fi fingerprinting, it allows quick localization. This also facilitates multi-floor navigation by using barometer readings to find stairways and elevators. In addition, ViNav helps smartphone users to search for and navigate to POIs with the information extracted from crowdsourced visual data. The new elements incorporated into ViNav, as shown in the field study, make it usable and well shaped. ViNav would be a good prim for any program that intends

to provide indoor navigation services based on the visual and sensor data provided by crowdsourcing.

REFERENCES

[1] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. M. Seitz, and R. Szeliski, "Building rome in a day," ACM Commune., vol. 54, no. 10, pp. 105–112, 2011.

[2] P. Bahl and V. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in Proc. 19th Anu. Joint Conf. IEEE Compute. Commune. Soc., 2000, pp. 775–784.

[3] L. Chen, L. Pei, H. Kuusniemi, Y. Chen, T. Kroger, and R. Chen,

"Bayesian fusion for indoor positioning using Bluetooth finger-prints," Wireless Pers. Commune., vol. 70, no. 4, pp. 1735–1745, 2013.

[4] J. C. Chow, "Multi-sensor integration for indoor 3D "PhD thesis, Univ. Calgary, Calgary, AB, 2014.

[5] 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. Serv., 2011, pp. 141–154.

[6] J. Dong, Y. Xiao, Z. Ou, and A. Yl€a-J€a€aski, "Utilizing internet photos for indoor mapping and localization–opportunities and challenges,"in Proc. IEEE Conf. Compute. Commune. Workshops, 2015, pp. 1–6.