**Vinav: A Vision-Based Indoor Navigation System For Smartphones**

Jiang Dong , Marius Noreikis, Yu Xiao, Antti Yla-J

**ABSTRACT**

Smartphone-based indoor navigation services are desperately needed in indoor environments. However, the adoption of them has been relatively slow, due to the lack of fine-grained and up-to-date indoor maps, or the potentially high deployment and maintenance cost of infrastructure-based indoor localization solutions. This work proposes *ViNav*, a scalable and cost-efficient system that implements indoor mapping, localization and navigation based on visual and inertial sensor data collected from smartphones. *ViNav* applies structure-from-motion (SfM) techniques to reconstruct 3D models of indoor environments from crowdsourced images, locates points of interest (POI) in 3D models, and compiles navigation meshes for path finding. *ViNav* implements image-based localization that identifies users’ positions and facing directions, and leverages this feature to calibrate dead-reckoning-based user trajectories and sensor fingerprints collected along the trajectories. The calibrated information is utilized for building more informative and accurate indoor maps, and lowering the response delay of localization requests. According to their experimental results in a university building and a supermarket, the system works properly and their indoor localization achieves competitive performance compared with traditional approaches: in a supermarket, *ViNav* locates users within 2 seconds, with a distance error less than 1 meter and a facing direction error less than 6 degrees.

**EVALUATION**

Authors implemented a prototype of *ViNav* to evaluate the feasibility of building a smartphone-based indoor navigation system using crowdsourced data.Authors have carried out field studies in different indoor environments including an office building and a supermarket. According to the field studies, the functionalities of *ViNav* have proved to work properly. Authors will focus on the performance evaluation of *ViNav*, and will introduce in this section the methodology and the experimental results, including a short summary of the feasibility study.

**INTRODUCTION**

Indoor navigation systems for smartphones are crucial in complex indoor environments such as airports, shopping malls and museums. Unfortunately, the adoption rate of in- door navigation systems is still very low, even though initial efforts into deploying them were taken several decades ago. There are multiple reasons for the slow progress. First of all, indoor navigation requires fine-grained and up-to-date indoor maps for calculating navigation routes and searching for points of interest (POI). Certain solutions1 ask users to provide photos of floor plans in public venues as their indoor maps. However, most of them lack details, or have not been kept up to date. Therefore, they can rarely be used directly for indoor navigation. Secondly, even in the case where accurate indoor maps do exist, most systems rely on pre-scanned radio maps or pre-installed hardware for localization, which are expensive to install and cumbersome to maintain. Thus, Authors saw it worthwhile to investigate whether Authors could develop an alternative indoor navigation method that would not need pre-created indoor maps, pre-scanned radio maps or pre-installed hardware.

Smartphones today are equipped with high-resolution cameras, and mobile users are willing to share some of their photos publicly, e.g. via photo-sharing websites like Flickr or Instagram. Furthermore, researchers have proven

*ViNav* also extracts POI information from the collected visual data and maps it into the 3D space, thus, users can easily locate POIs.

With SfM-based 3D models, the system provides image- based localization that identifies user’s position and facing direction using the photos taken in situ. The problem with this approach is that the response time increases as the sizes of the 3D models increase, because location is determined by matching 2D features in a query photo against 3D points contained in the 3D models. To enable fast localization, *ViNav* employs model partitioning based on the density of 3D points, and selects partitions for feature matching based on Wi-Fi fingerprinting. Experimental results demonstrate decreased localization response delays for this partitioning scheme even for large 3D models. By utilizing the image- based localization, the navigation mesh and the identified POI position, The system can provide a navigation path and guide mobile users to their destinations.

In this paper Authors present the design, implementation and thorough evaluation of *ViNav*. The contributions of this paper are summarized as follows:

* Authors address several technical challenges to enable *ViNav*: Authors combine the pedestrian paths recognized from crowdsourced user trajectories to complement the crowdsourced 3D models of indoor environments; Authors fuse barometric pressure data with user trajectories to detect connecting paths between floors; Authors extract POI information from crowdsourced visual data and map the extracted POI into the 3D space; Authors speed up the localization process by density-based model partitioning and fingerprint-based partition selection.
* Authors implement a prototype of *ViNav* to evaluate its feasibility and performance. Experimental results demon- strate that *ViNav* works properly in multi-storey build- ings and shows good performance. According to their experiments carried out in two different indoor envi- ronments, *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.

**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 ,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 per- formance evaluation of *ViNav*, Authors measure the following metrics:the accuracy of the indoor layout extracted from the 3D models, the accuracy of detecting stairs and elevators in multi-floor buildings, the performance of POI detection, and the performance of the 3D-model- based indoor localization.

**SUMMARY**

From the experiments and analysis presented in this research paper, Authors reach to the conclusion that it is feasible to build a well performing 3D-model-based indoor navigation system, using only the photos and sensor data collected from smartphones. The quality of 3D models, nevertheless, is affected. VisualSFM does not support pre-loading models since it is a closed source software.

the dead reckoning approach, Nguyen et al. utilized active learning for automatically identifying strategically important locations which should be labelled manually. TraviNavi employed traces of fingerprints recorded from entrance POIs to be used for leader-follower navigation. Rather than manually labelling Wi-Fi fingerprints, ViNav utilizes arbitrary photos taken along user trajectories for automatic calibration, utilizing the feature of image-based localization. Furthermore, ViNav provides navigation to arbitrary destinations with no requirement of following a prerecorded trace. Image-based localization systems allow users to locate themselves by simply taking photos from where they are. State-of-the-art systems such as **Travi-Navi employs image histogram matching**, Liu et.al utilized deeplearning approach for matching and tracking, while Huang et al. proposed image feature matching for panoramic images to obtain user’s position. However, this process of image feature matching is slow due to heavy computation. To accelerate this process, Lu et al. proposed to use the visual words and the approximate nearest neighbour methods. Similarly, Sattler et al. proposed a framework for efficient 2D-to-3D matching based on visual vocabulary quantization and a prioritized correspondence search. Sextant utilizes physical features (i.e., logos or paintings) as reference objects to measure user positions. The challenges of designing such a POI-based localization system include identifying distinguishable POIs, obtaining accurate positions of POIs, and etc. ViNav utilizes features extracted from images as references to locate users. Distinguishable features are more widely available than distinguishable POIs in different indoor environments. ViNav further utilizes Wi-Fi fingerprints for reducing the search space and minimizing response latency while utilizing SfM based pose calculation to provide high accuracy. In addition to location, a walking direction of a user is also important for creating and updating navigation instructions. WalkCompass detects the walking direction within a few steps using sensors available on smartphones, while Husen et al. employed a dense Wi-Fi infrastructure to determine the orientation. In their work, user’s facing direction is considered to be equal to the facing direction of the smartphone camera, and is obtained from the camera pose provided by image-based localization.

**CONCLUSION**

*ViNav*, a low-cost and whole-system solution for indoor navigation in this paper. *ViNav* is partially built on top of several existing techniques, e.g., SfM and fingerprinting., *ViNav* brings new functions and addresses several technical challenges to enable mobile crowdsensing-based indoor navigation. It utilizes crowd- sourced visual data to build 3D models of indoor spaces of interest, and detects pedestrian paths from crowdsourced user trajectories. It enables fast localization by taking ad- vantage of the high hit-rate and low response delay of Wi- Fi fingerprinting. It also supports multi-floor navigation by locating stairs and elevators with the help of barometer readings. Moreover, with the information extracted from crowdsourced visual data, *ViNav* enables mobile users to search for POIs and navigate to them.