Indoor localization and navigation using smartphones augmented reality and inertial tracking

Conference Paper · December 2013

DOI: 10.1109/ICECS.2013.6815564

4 authors, including:

Buti Mohammed Al Delail
Khalifa University of Science Technology & Res…

Luis Weruaga

Jamal Zemerly
Khalifa University of Science Technology & Res…

Indoor localization and navigation using smartphones augmented reality and inertial tracking

Buti Al Delail, Luis Weruaga, M. Jamal Zemerly
Electrical and Computer Engineering Department
Khalifa University of Science, Technology and Research
United Arab Emirates
{buti.aldelail, luis.weruaga, jamal.zemerly}@kustar.ac.ae

Jason W.P. Ng
Etisalat BT Innovation Center
BT Innovate & Design
Abu Dhabi, UAE
jason.ng@bt.com

***Abstract*—Over the last years, indoor localization and navigation is becoming a hot topic. With the increasing number of buildings, indoor positioning and navigation has turned out to be more important than outdoors. In the literature, many papers discuss wireless based indoor positioning systems. Essentially based on Wireless Fidelity (Wi-Fi), Bluetooth, Radio Frequency Identification (RFID) or existing solutions that imply the measurement of radio signals. In this paper, we evaluate an indoor image-based positioning system that takes advantage of smartphones augmented reality (AR) and inertial tracking. The excellent computing capabilities in todays highend phones or smartphones in combination with its resources of sensors, such as Global Positioning System (GPS), inertial sensors, camera, wireless receivers, are powering the mobile application sector to the extent of becoming the fastest growing one in data communication technologies. AR as an emerging technology has the potential of creating new types of indoor location based services for the near future. Here, we show some of the AR capabilities combined with inertial tracking for localization and navigation.**

***Index Terms*—Indoor Localization, Augmented Reality, Inertial Tracking, Indoor Navigation**

I. Introduction

With the rapid advancements in smartphones technology, the ability of obtaining the position of devices and persons is the driving core for many applications. GPS is now available in every smartphone and being used for all kinds of applications, not only by the traditional maps application. Nowadays the location used for calculating accurate local time, current weather data or providing customized information that is useful to individual user. Also, it is the best existing solution for outdoor navigation with a precision of up to 1 meter [1]. However, in indoor environments, the radio signal quickly weakens because of the roof thickness, walls and nearby objects [2]. Knowledge about the orientation and location delivered by current handheld devices allows for context-aware information services tailored to individual user according to their preference. In that regard, the excellent performance in mobile visual computing [3] permits the detection and tracking of markers and planar objects in real time. This information can be used for mobile visual location recognition [4] in dense outdoor or urban areas where GPS measurements are either unavailable or inaccurate.

On the other hand, localization in indoor scenarios is dominated by wireless-based positioning techniques in combination with inertial sensors, and others [5], [6]. The combination of computer vision and inertial measurements has been also explored in indoor environments [7], but this requires uncomfortable wearable computing equipment. A breakthrough application derived somewhat from mobile visual computing is augmented reality (AR), namely, the display of virtual 3D objects that merge seamlessly with the actual video scene captured by the device camera. As AR is founded on detection of a known marker, as well as its tilt in 3D coordinates [8], when the marker is in a permanent static location, de facto localization is accomplished.

II. Concept and Related Work

The work in [9] the first step towards mobile and wearable augmented reality. The hardware design has many disadvantages such as the difficulty to carry-with and unlikely to be used by everyday users. After that, the hand-held augmented reality came in the form of PDAs and laptops with frontally mounted camera, then when smartphones emerged with enhanced computational power it became a convenient hardware platform for AR systems, primarily because of the ease of use for inexperienced users [10].

Following the previous concept, this article proposes a scalable mobile based system for indoor location detection and tracking using a combination of image marker recognition and inertial measurements. This localization service is core to deliver a context-aware information system built around an augmented reality software layer. The AR layer informs the user of a nearby point of interest, overlaying self-explanatory 3D virtual objects related to the location on the real-time video capture. The points of interest nearby also shown in a 360 degree fashion supported on the device compass readings. Any AR virtual object is “clickable” or “touchable”, such that the user obtains information related thereupon. Recently there has been similar work related to vision-based indoor localization, navigation and the use of augmented reality. For instance, authors in [11] proposed a vision-based mobile indoor localization and navigation system with Augmented or Virtual Reality (VR) interfaces. In this paper, we use AR only since the visual localization and inertial tracking is fairly accurate for indoor use, insuring that the graphical contents will augment the video stream as intended. Not only that, but the research from the same authors showed that users preferred the AR interface over VR interface. Therefore, the system [12] was initially developed for the University campus where the authors work. Another use for this application is the work in [13], which leverages the smartphones AR capabilities for a treasure hunt games. Also, the work in [14] using Near Field Communication (NFC) tags and Quick Response (QR) Codes for indoor localization and map guidance.

III. Vision Based Localization

Using vision as a base for localization has a number of advantages over other real-time localization methods. It does not rely on external electronic devices (e.g. Bluetooth or WiFi access points) and takes advantage of the camera and enough processing power, which put together in every smartphone. Vision-based tracking is currently the most accurate type of tracking, which is measured in pixels. A lot of progress has been done in the visual tracking area, where approaches are gradually moving from marker-based tracking to Simultaneous Localization And Mapping techniques (SLAM) [15] [16]. Using markers is one way of telling the computer how, what and where to augment the real-world view. Many visual techniques are based on tracking high-level image key-points (such as Speeded up Robust Feature (SURF), Binary Robust Invariant Scalable Keypoints (BRISK) and Fast Retina Keypoint (FREAK) [17]) instead of pixels. These key-points are not just used for tracking, but also for object and image detection for embedded devices with low memory and computational power, enabling faster object recognition and AR registration in low-end devices.

In order to get the initial location, the smartphone user must catch sight of an image marker using the camera. Each marker has a permanent indoor location that is stored in the database. When a marker is detected (see Fig. 1 for an illustrative explanation of the feature or fiducial points of a certain marker), its unique identifier is used to search the database for the location of that marker. Since the user is near to the marker, this can be used as the user location. Furthermore, the visual tracking can be combined with the inertial sensors (compass, accelerometer and gyroscope) to get better tracking abilities, some promising results in visual-inertial sensor fusion can be seen in [18].

The challenge here is that vision-based localization heavily relies on the recognition of known images in the environment, and when the environment is large, the number of images increase. The smartphone can not handle processing a large database of images. A feasible solution is moving the image recognition processing to a remote workstation. Cloud-based recognition (available in Vuforia [19]) enables the smartphone to recognize a vast number of images in real-time. Additionally, positioning the images database on the cloud allows easy update in case changes occurred in the environment.

IV. Augmented Reality Interface

The concept of AR makes it usable for unlimited applications, the advantage here is that the AR interface naturally fits with vision-based localization. At the same time that the image is being tracked and the users position is known, an informative computer generated graphical objects augment parts of the screen. Fig. 2 shows a snapshot of the 3D AR objects superimposed on the original view captured by the camera. The acquired facts of the detected marker (shown in Fig. 1) delivered by Vuforia corresponds to the distance and tilt of the scene is used to render the 3D object in the "correct" place. Moreover, "touchable" virtual buttons, provided also by the Vuforia SDK [19], allow for visual interaction with the identified image, and used for information retrieval. The system becomes portable and easily configurable by allowing 3D object data to be downloadable from the cloud.

Fig. 1. Target image features and fiducial points.

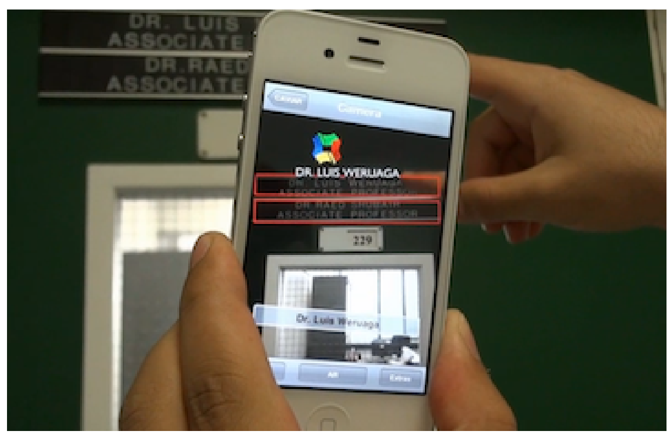

Fig. 2. Augmented reality: 3D virtual object displayed on the video scene; 'clickable" virtual buttons (in red line); when a virtual button is pressed, the related information (faculty name) is displayed.

V. Inertial Tracking

Inertial tracking overrides the visual SLAM described in Section III when the tracked image goes out of frame. There are many inertial tracking systems that detect the user footsteps in order to perform SLAM [20], but only recently, these techniques are being explored to be used in smartphones. Despite the challenge in dealing with measurement errors and dead reckoning from the accelerometer, recent methods in signal recognition, position prediction and correction looks to be promising. Therefore, the availability of inertial sensors and processing power in smartphones, and being the most common device that people carry, it is the most favorable hardware for inertial tracking.

The method we use detects the user footsteps based on the accelerometer readings (see Fig. 3), and updates the user location 2 ft towards the direction read from the compass. This implementation is reasonably accurate, provided that the user knows how to use it.

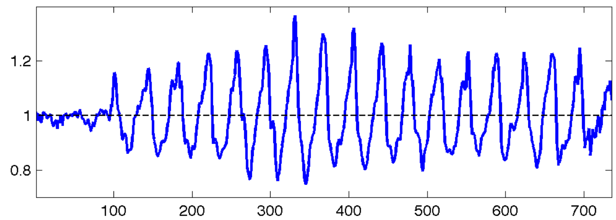

Fig. 3. Magnitude of the accelerometer readings with walking user: every cycle in the signal corresponds to a user footstep.

The result from the initial approach is shown in Fig. 4, where the walking path is more than 300 ft long, resulted in localization error of less than 3 meters at the destination. However, it can be seen that the overall error at certain points is more. Furthermore, the compass heading is not very reliable, and using that only with no calibrations may produce errors as high as 90°. Therefore, to improve the accuracy we calibrate the compass and add contraints on the estimations to insure that the estimated location is within the path. The error correction mechanism uses the compass direction to produce a field of view, where the estimated heading is considered towards the nearest node within view, thus, not relying on the exact heading degree for the direction. The improvement can be seen in Table I for the same path, the heading error with the correction mechanism depends on the user location, walking path and the node infront.

TABLE I
INERTIAL TRACKING WITH AND WITHOUT CORRECTION

Comparison	No Correction	With Correction
Maxiumum Heading Error	+-5°	+-2.3°
Final Position Accuracy	3 meter	1 meter
Walking Path Length	306.7 meter	309.4 meter

Further efforts are being put to track the location using inertial sensors, disregarding the need of having the device held in a fixed position. Fig. 5 shows a sample of the accelerometer reading captured while the user is walking, with the device is being inside the pocket. Principal Component Analysis (PCA) is put to use to identify the direction of motion. The pricipal component is calculated for a window size of 1 second (+-0.2 sec.) to remove the signal noise extract from the instant acceleration in order to obtain a high-level vector of the gravity and motion.

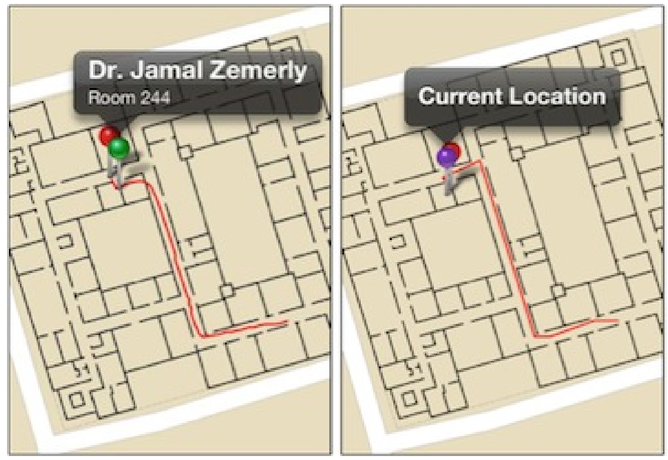

Fig. 4. Trace (red) obtained from the inertial location tracking module for a walking path length of about 300 ft. The left snapshot shows the original estimated walking path, in contrast, the right snapshot shows the corrected path.

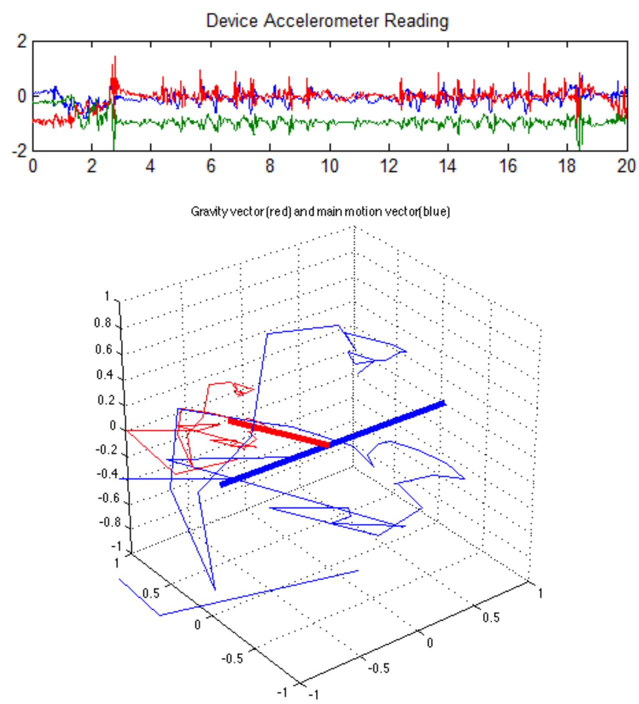

Fig. 5. Magnitude of the accelerometer readings and the direction of motion

VI. Indoor Navigation

The localization system can be used efficiently to obtain the user location and search for other places. Determining the path from the user location to the destination is the key to providing indoor navigation. In order for the system to search for the path, intersections within the building are added to the location database with links that identify which locations can be directly connected. This data can be used to search for the shortest distance path between any two locations.

Therefore, we use the A* algorithm [21] to find the shortest path. The algorithm ensures finding the shortest path in the fastest possible way without the need of exhaustive visiting or testing of all locations.

Moreover, navigation is not about finding areas and paths, but also guiding and monitoring the movement of the user. Here, where AR plays another role in the interface. Instead of looking at a 2D map while navigating, AR can be used to show virtual guiding objects to the user view. The previously described SLAM allows computer graphics (e.g. signs and directions) to be mapping and updated simultaneously to the camera view. For instance, points of interest AR displayed in Fig. 6. This method relies on the compass together with the gyroscope, to produce a 360-degree scene with the points of interest (PoI) in the vicinity of the current mobile user location. As proof of concept and for the sake of simplicity, the PoIs are presented with a simple clickable button.

Fig. 6. Indoor 360-degree points-of-interest view.

VII. Conclusion and Future Work

This paper presents a new augmented reality system method for indoor localization and navigation, thus allowing users to be aware of their locations and making it easier to find places with reasonable accuracy. The use of image recognition indoors enables the system to run on any campus, provided the database for recognizing markers, obtaining the location and displaying the information in AR exists. Also, results show that inertial navigation provides a suitable approach to track the user location indoors.

Moreover, smartphone based augmented reality separates Augmented Reality from Wearable Computing as initial AR systems hardware consisted of a Head-mounted Display (HMD), which must be wearable. Therefore, wearable AR glasses have been a research and development project by major companies (such as Googles Glass Project [22]), which will not just be the next-generation mobile AR hardware platform, but also replace smartphones. Although there is a gap in technology to enable AR to be a part of peoples daily life, there is no doubt that AR is the future. The system in this paper can be deployed on any hardware with the sufficient resource and sensors available in any modern smartphone.

References

[1] H. Koyuncu, S. H. Yang, "A Survey of Indoor Positioning and Object Locating Systems," *IJCSNS International Journal of Computer Science and Network Security*, Vol. 10, No. 5, pp. 121-128, 2010

[2] E. Coca, V. Popa, and G. Buta, "An Indoor Location System Performance Evaluation and Electromagnetic Measurements," *IJCSNS International Journal of Computer Science and Network Security*, Vol. 10, No. 5, pp. 121-128, 2010

[3] B. Girod, et al., "Mobile visual search," *IEEE Signal Processing Mag.*, vol. 28, no. 4, pp. 61–76, Jul. 2011.

[4] G. Schroth, et al., "Mobile visual location recognition," *IEEE Signal Processing Mag.*, vol. 28, no. 4, pp. 77–89, Jul. 2011.

[5] L. Klingbeil, et al., "A modular and mobile system for indoor localization," *Intl. Conf. Indoor Positioning & Navigation*, 2010, pp. 1–10.

[6] S. Lee, B. Kim, H. Kim, R. Ha, and H. Cha, "Inertial sensor-based indoor pedestrian localization with minimum 802.15.4a configuration," *IEEE Trans. Industrial Informatics*, vol. 7, no. 3, pp. 455–466, Aug. 2011.

[7] D. Chdid, R. Oueis, H. Khoury, D. Asmar, and I. Elhajj, "Inertial-vision sensor fusion for pedestrian localization," *IEEE Intl. Conf Robotics Biomimetics*, 2011, pp. 1695–1701.

[8] D. Wagner, et al., "Real-time detection and tracking for augmented reality on mobile phones," *IEEE Trans. Visual. Comp. Graphics*, vol. 16, no. 3, pp. 355–368, May/Jun. 2010.

[9] S. Feiner, B. MacIntyre, and T. Hollerer. Wearing It Out: First Steps Toward Mobile Augmented Reality Systems. *In International Symposium on Mixed Reality (ISMR 99)*, pp. 363-377, 1999.

[10] D. Schmalstieg and D. Wagner, "Experiences with Handheld Augmented Reality," *Proc. Sixth IEEE and ACM Intl Symp. Mixed and Augmented Reality*, 2007.

[11] A. Mller, M. Kranz, R. Huitl, S. Diewald, L. Roalter, "A mobile indoor navigation system interface adapted to vision-based localization," *11th International Conference on Mobile and Ubiquitous Multimedia*, ACM, pp.4-14, 2012.

[12] B. Al Delail, L. Weruaga, M. J. Zemerly, "CAViAR: Context Aware Visual indoor Augmented Reality for a University Campus," *IEEE Intl. Conf on Web Intelligence and Intelligent Agent Technology*, pp.286-290, Dec. 2012.

[13] Z. Balint, B. Kiss, B. Magyari, K. Simon, "Augmented Reality and Image Recognition Based Framework for Treasure Hunt Games," *10th Jubilee International Symposium on Intelligent Systems and Informatics*, 2012, pp. 148–152.

[14] O. Al Hammadi, A. Al Hebsi, M. J. Zemerly, J. W. P. Ng, "Indoor Localization and Guidance Using Portable Smartphones" *IEEE Intl. Conf on Web Intelligence and Intelligent Agent Technology*, pp.337-341, Dec. 2012.

[15] D. Stricker, G. Bleser, "From Interactive to Adaptive Augmented Reality," Ubiquitous Virtual Reality (ISUVR), 2012 International Symposium, pp.18-21, 22-25 Aug. 2012.

[16] G. Klein, D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium, pp.225-234, 13-16 Nov. 2007.

[17] A. Alahi, R. Ortiz, P. Vandergheynst, "FREAK: Fast Retina Keypoint," Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference, pp.510-517, 16-21 June 2012.

[18] G. Bleser, D. Stricker, "Advanced tracking through efficient image processing and visual-inertial sensor fusion," *IEEE Virtual Reality 2008*, IEEE, pp.137-144, 2008.

[19] Vuforia SDK, Qualcomm Inc., https://developer.vuforia.com, Apr. 2013.

[20] P. Robertson, M. Angermann, B Krach, "Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors", *In 2009 Proceedings of the 11th international conference on Ubiquitous computing (Ubicomp '09).* ACM, New York, NY, USA, pp.93-96.

[21] P. Hart, N. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost path," *IEEE Transactions of Systems Science and Cybernetics*, pp.100-107, 1968.

[22] "Google gets in your face [2013 Tech To Watch]," *Spectrum, IEEE*, vol.50, no.1, pp.26-29, 2013.