Accurate indoor navigation using Smartphone, Bluetooth Low Energy and Visual Tags

Gaetano Carmelo La Delfa, Vincenzo Catania

Department of Electrical, Electronics and Computer Engineering (DIEEI) University of Catania, Italy
{Vincenzo.Catania,gladelfa}@dieei.unict.it

Abstract. Every moment of our day takes place in a location, identified by geographical coordinates which connect physical and digital world. For that reason it's important to know the user position: it enables new location-related services and give to the people the opportunity to find their way in complex environments such as hospitals. The best approach to realize an indoor navigation system is to exploit the potential of the smartphone in some way. Different techniques were proposed in scientific literature, but the topic is still matter of research. In this paper we propose a smartphone-based indoor navigation system which use both Bluetooth Low Energy(BLE) and a visual-tags system. A BLE beacons system is deployed in the environment to segment a big area into smaller areas and to perform a rough background localization. An accurate indoor navigation is performed using the camera to decode a visual-tags system deployed onto the room's floor.

Keywords: Indoor navigation, Visual Tags, ArUco, Bluetooth Low Energy, Smartphone

1 Introduction

The massive diffusion of smartphone, which is a perfect gateway between real and digital life, enables a lot of services for the user. Most of them are strictly related to the position so it is important to know where the user is in a certain moment and how he can navigate towards a destination. Thinking to the healthcare sector, it is common that patients and visitors, even if there are signage deployed on the environment, lost their way inside large hospitals. Moreover, patients with physical handicaps can experiment more difficulties than other people to orient themselves inside the building. An accurate smartphone indoor navigation system can be very useful for all these specific situations: it is possible to provide both visual and voice turn-by-turn navigation for people with visual or hearing deficiency, create specific POI and provide a way to reach them, find a doctor or locate a patient's room in a medical facilities, and lot more. While the GPS Technology has an excellent performance outdoor, it fails into the indoor environments, due to the fact that the GPS signal is not available inside the buildings. A lot of indoor navigation techniques were proposed in literature, with different accuracy and complexity levels, but there is not a unique solution today, and the problem is still matter of research.

In this paper we suggest a novel approach which aims at solving the problem by performs indoor localization/navigation using only BLE and the embedded camera on the smartphone. Our approach is based on the following simple ideas/hypothesis:

1. When the user has the navigation app in background mode, he does not need to navigate inside the building, and probably does not need to know his position with an high accuracy level. A low level accuracy position estimation using the BLE localization system can give a good information on the user position.
2. When the user needs high accuracy, he launches the navigation app. Thanks to the fact that he has the device on the palm of his hand, necessarily the smartphone's camera will be directed towards some part of the floor. A visual-tags system properly deployed onto the floor (at known points) and an algorithm capable to detect/decode in real time these tags, can estimate the position of the user with an high level of accuracy (depending on the density of tags) and let him to navigate towards the target.
3. From the point of view of the person who install the indoor navigation system, it is simple to print the visual tags and to deploy them onto the floor, using a web application to generate the tags and to determine the best position where the tags should be installed.

The remainder of this paper is organized as follows: Section 2 discuss about the main techniques actually used for indoor navigation. Section 3 discuss how to partition an area, identified by a single BLE beacon, in microareas, and how to assign the visual tag which identifies each microarea. Section 4 discusses about some of the most important visual-tag systems, and focuses on the visual-tags system chosen in this paper. Finally, in Section 5 we make some conclusions and talk about future works.

2 Related Works

In last years, various approaches for accurate indoor localization/navigation have been proposed by researchers from all over the world [12]. In the following we give a brief and non exhaustive summary of these approaches, and evidencing some drawbacks:

- *Proximity approach*: the nearest network node is determined using some kind of signal emitted by the node. The smartphone can sense this signal and determine its own position related to the known network node position [9]. This approach is simple and low cost but provides low-level accuracy.
- *RSSI (Received Signal Strength Indication) fingerprint*: the smartphone measures the RSS values in the environment and compares them with previously measured values, which are saved in a database (the fingerprint), to stimate the position of the user and to perform indoor navigation. This approach, often used with wi-fi signals which are available for free in the buildings, provides a good level of accuracy but needs experienced personnel to create the fingerprint for each new area [10], which bring to high cost and non-scalable solutions that need frequent reconfigurations.
- *Dead Reckoning*: the indoor navigation is performed by using the inertial sensors embedded on the device. Usually, due to the unacceptable drift error of the inertial sensors, a step-counter system is used to estimate the distance covered by the user. This approach is easy to deploy and low cost but needs a periodic position recalibration to reset the error which quickly becomes too high for indoor environments [11],[5].
- *Visual tag approach*: a visual-tags system is deployed inside the building at known points. The user, using the embedded camera on the smartphone, can navigate inside the building [1]. This approach provides a variable level of accuracy (depending on density of tags), but has high latency due to tag detection and is uncomfortable for the user who has to find the tag, and focus the camera on it.
- *hybrid approach*: usually multiple techniques are used together to improve the efficiency of the localization/navigation [13],[2].

3 Room segmentation

In the proposed approach, the whole target building is segmented into *Room Level Areas* identified by BLE beacons, installed in a known positions. The segmentation is useful for two reasons: (1) the smartphone can sense a BLE beacon (or more of them) to determine in which Area the user is even when the app is in background, and get notified when he moves inside a specific Area. (2) The complexity of the visual-tags system is reduced because it is possible to reuse the same set of tags for different Areas identified by different BLE beacons.

Each *Room Level Area* is divided into a number of *Micro Level Areas*, according to the *Voronoi-based Segmentation*. This is a mathematical way of partitioning a space into a number of regions. If we consider a plane π and a finite set of points S, the Voronoi segmentation of π is the partition of the plane which associate a region $V(p)$ to every $p \in$ S, in a way that all the points of $V(p)$ are closer to p than every other point in S [6]. Fig. 1 show an example of Voronoi segmentation of a plane. We use the Voronoi segmentation to automate the partitioning of the *Room Level Area* into *Micro Level Areas*, which otherwise must be manually done. To generate the Voronoi segmentation,

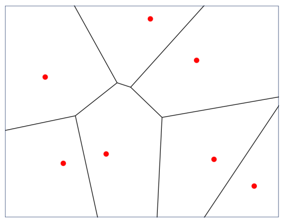
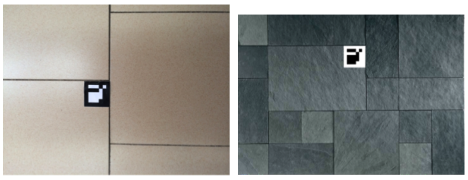


Fig. 1. Voronoi segmentation of a plane.

a set of points p is distributed in a raw map of the *Room Level Area*. Each point has an associated visual tag, which identifies the correspondent Voronoi region $V(p)$ in the *Room Level Area*. It is possible to concentrate points in some areas, and avoid to set points in non-accessible areas. The minimum distance between tags depend on the used camera and on the desired accuracy. To limit the number of tags deployed onto the floor and to reduce the probability to have multiple tags in the same captured frame, a minimum distance of $1,5\ m$ between each tag must be set (this is an experimental value that need further investigations). To make quicker and more efficient the tag detection, we choose a tag color chromatically far from the color pattern of the floor. The Fig. 2 show two examples of floor patterns with associated tags, to better explain the concept: The tag's border is used by the detection algorithm to detect the tag on the frame captured by the camera. As we can see, the tag placed in a light floor color's pattern has a black border, while the tag placed in a dark color's pattern, has a white border. These colors guarantee high contrast between floor's color pattern and tag's border in both light and dark cases, and greatly improve the speed and the efficiency of the detection algorithm.



(a) *Light* (b) *Dark*

Fig. 2. ArUco tag deployed on a light 2(a) and dark 2(b) floor pattern

Once all the beacons have been installed, all the points are set and all the tags are generated and deployed on the floor, the room segmentation is ended. These data are uploaded on the smartphone to use them for indoor localization/navigation.

4 Visual tags

Visual tags are small, machine-readable images, used to codify the information. Actually there are a lot of visual tag systems, which have different features. For the purposes of this paper, we can split these systems in two big categories: (1) visual tags used to store a large amount of data such text and URL and (2) visual tags used to store a small set of symbols. In the following we briefly talk about the most famous visual tag belonging to the first category (QRCode), and about two of the many visual tags belonging to the second category (Vuforia marker, ArUco tag):

- *QRCode*: it is a popular 2D barcode invented in 1994 by Denso Wave [4] and used to store a large amount of data. It uses the reed-solomon error correction algorithm to manage the error, and can be easily decoded by any smartphone. It is rotation-independent, has a strong tolerance to deformations and partial occlusion and it is free of any license.
- *Vuforia Frame Marker*: It is a visual tag realized by Qualcomm Technologies. It can be detected by any smartphone using the Vuforia SDK [8]. It can store up to 512 symbols and it has a low decoding latency. It is rotation-independent, has a strong tolerance to deformations, it is free to use but not opensource, so it is not possible to modify the source code.
- *ArUco Tag*: It is a visual tag realized by the A.V.A. group from University of Cordoba. It can be detected by any smartphone using the ArUco library, a minimal openCV based library. ArUco library let to generate configurable dictionaries of tags (both in size and number of bits) and maximize the inter-tag distance and the number of bit transitions. The detection/decoding algorithm has a low latency and strong tolerance to deformations. Is rotation-independent, cross-platforms (thanks to the fact that the library is based on openCV exclusively), opensource (BSD license) and embed an error correction routine. [7].

To perform a real-time indoor navigation as described in this paper, we need a visual-tags system which is robust to rotations and deformations, and which is as fast as possible in detection/decoding phases. Moreover, this visual-tags system must be opensource to allow us to modify the general purpose detection algorithm on which it is based in order to adapt it to our scenario: (1) tags placed on the floor, which exhibit a regular and uniform color pattern (as we saw before in Fig. 2(a) and Fig. 2(b)). (2) Almost constant size of the tag inside each captured frame, due to the fact that the distance between smartphone camera and floor is more or less constant (we can exploit these features to improve the speed of the tag detection process).

In our analysis we excluded the Vuforia marker, due to the fact that it is not opensource, and QRCode, due to the fact that because it stores a large amount of data, the detection/decoding performances are not suitable for our purposes. The visual tag which best fits our requirements is ArUco.

(a) *QRCode* (b) *Vuforia* (c) *ArUco*

Fig. 3. Examples of visual tags

To generate the personalized dictionary of tags, ArUco algorithm start from an empty dictionary D which is incrementally populated with new tags selected (from the space of all possible tags) by using a stochastic process that assign, for each iteration, more probability to tags with an higher number of bit transitions. A inter-tag distance threshold t is initially set: if the distance between the generated tag and these in D is greater than t then the tag is added, otherwise is rejected and a new tag is randomly selected. To guarantee the convergence of the algorithm, the distance threshold t is initially set to the maximum value that the dictionary can have, then is reduced after a certain number of unproductive iterations. The process stop when the desired number of tags is reached.

To detect the ArUco tags, the algorithm follow the process briefly showed on Fig. 4: The original image *(a)* is converted to grayscale and segmented *(b)* using a local adaptative thresholding approach. Then a contour extraction *(c)* and a filtering process to discard the image contours which are irrelevant for the detection *(d)* are performed. The perspective projection is removed by computing the homography matrix and the result is thresholded to obtain the black and white ArUco tag *(e)*. The next step is to divide into a regular grid the resulting image and assign the value 0 or 1 to each element of the grid *(f)* (The black border around the tag is used by the algorithm to detect the tag on the frame captured by the camera). Once the code

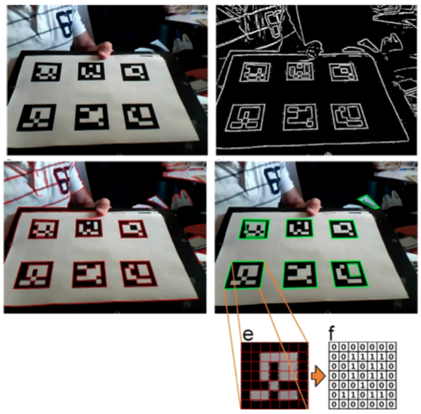


Fig. 4. AruCo tag detection

of a tag is extracted, the algorithm check if it is included in the previously generated dictionary (so, it is a valid tag): if no match is found, it apply the error correction routine. Moreover, the algorithm estimate the pose of the tag respect to the camera.

To perform an almost real-time detection/decoding we need a tag's size of $7cmx7cm$ or more. Latest generation smartphone and good light conditions, as well as detection/decoding algorithm enhancements, will improve the accuracy of the approach and allow to choose a smaller size for the tags.

5 Conclusions and future works

The goal of our work is from one side: (1) to give a methodology to enhance the process of self-localization/navigation, (2) to reduce the undesirable cognitive workload for the user to self-locating himself, (3) to merge the self-localization/navigation technique with BLE localization. We achieve all of these by deploying a visual-tags system on the floor (based on the idea that when the user launches the application to navigate inside the building, his camera is necessarily directed towards the floor), and by using the rough BLE localization when the app is in background mode, according to the analysis of the users behavior in the most common uses cases. From the other side our goal is to present a scalable, low cost, easy to deploy for everyone and easy to modify indoor localization/navigation infrastructure. The approach can be successfully used in a lot of applicative scenarios such as in the hospitals, in the airports, in the museum, etc. We are planning to modify the ArUco algorithm to improve the tag's detection/decoding speed by exploiting two features: (a) uniform pattern of the floor; (b) fixed size of the tag inside the captured frame. We are also planning to integrate a dead reckoning algorithm to track the user between two tags and investigating the opportunity to embed the tags into the tiles.

References

1. Bellot Arias, S.: Visual tag recognition for indoor positioning (2011)
2. Bissig, P., Wattenhofer, R., Welten, S.: A Pocket Guide to Indoor Mapping. In: Workshop on Positioning, Navigation and Communication (WPNC), Dresden, Germany (March 2013)
3. Chandgadkar, A., Knottenbelt, W.: An indoor navigation system for smartphones. Imperial College London (2013)
4. Denso Wave: QR-Code Standard (2010), `http://www.denso-wave.com/qrcode/qrstandard-e.html`
5. Emilsson, A.: Indoor navigation using an iphone (2010)
6. Gallier, J.: Notes on Convex Sets, Polytopes, Polyhedra Combinatorial Topology, Voronoi Diagrams and Delaunay Triangulations. Rapport de recherche RR-6379, INRIA (2007), `http://hal.inria.fr/inria-00193831`
7. Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F.J., Marín-Jiménez, M.J.: Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition 47(6), 2280–2292 (2014)
8. Incorporated, Q.: Qualcomm vuforia (2014), `https://developer.vuforia.com/`
9. Liu, H., Darabi, H., Banerjee, P., Liu, J.: Survey of wireless indoor positioning techniques and systems. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 37(6), 1067–1080 (2007)
10. Liu, Y., Wang, Q., Liu, J., Wark, T.: Mcmc-based indoor localization with a smart phone and sparse wifi access points. Pervasive Computing and Communications Workshops, IEEE International Conference on 0, 247–252 (2012)
11. Liu, Y., Dashti, M., Zhang, J.: Indoor localization on mobile phone platforms using embedded inertial sensors. In: WPNC. pp. 1–5. IEEE (2013)
12. Schwartz, T.: The Always Best Positioned Paradigm for Mobile Indoor Applications. Phd-thesis, Saarland University, Saarbrcken (3 2012)
13. Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M., Choudhury, R.R.: No need to war-drive: Unsupervised indoor localization. In: Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services. pp. 197–210. MobiSys '12, ACM, New York, NY, USA (2012), `http://doi.acm.org/10.1145/2307636.2307655`