



Heriot-Watt University
Research Gateway

Indoor Positioning System Using Depth Maps and Wireless Networks

Citation for published version:

Duque-Domingo, J, Cerrada, C, Valero, E & Cerrada, JA 2016, 'Indoor Positioning System Using Depth Maps and Wireless Networks', *Journal of Sensors*, vol. 2016, 2107872.
<https://doi.org/10.1155/2016/2107872>

Digital Object Identifier (DOI):

[10.1155/2016/2107872](https://doi.org/10.1155/2016/2107872)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

Published In:

Journal of Sensors

Publisher Rights Statement:

Copyright © 2016 Jaime Duque Domingo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Article

Indoor Positioning System Using Depth Maps and Wireless Networks

Jaime Duque Domingo ^{1*}, Carlos Cerrada ¹, Enrique Valero ² and J. A. Cerrada ¹

¹ Departamento de Ingeniería de Software y Sistemas Informáticos, ETSI Informática, UNED, C/Juan del Rosal, 16. 28040 Madrid, Spain

² School of Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

* Author to whom correspondence should be addressed; jaimeduque@amenofis.com

Abstract: This work presents a new *Indoor positioning system* (IPS) based on the combination of *WiFi Positioning System* (WPS) and *depth maps*, for estimating the location of people. The combination of both technologies improves the efficiency of existing methods, based uniquely on wireless positioning techniques. While other positioning systems force users to wear special devices, the system proposed in this paper just requires the use of *smart-phones*, besides the installation of RGB-D sensors in the sensing area. Furthermore, the system is not intrusive, being not necessary to know people's identity. The paper exposes the method developed for putting together and exploiting both types of sensory information with positioning purposes: the measurements of the level of the signal received from different access points (APs) of the wireless network, and the *depth maps* provided by the RGB-D cameras. The obtained results show a significant improvement in terms of positioning with respect to common WiFi-based systems.

Keywords: positioning; WPS; RGB-D sensors; Kinect; WiFi; *fingerprint*; trajectory; skeletons; *depth map*

1. Introduction

Indoor Positioning Systems (IPS) are techniques used to obtain the position of people or objects inside a building [1]. Among these, *WiFi Positioning Systems* (WPS) [2] are those based on portable devices, such as *cell phones*, to locate people or objects by means of the measurements of the level of the signal received from different access points (APs), i.e. WiFi routers.

In the field of people and objects detection, other technologies, such as those based on artificial vision, have been increasingly used. In fact, object recognition can be considered as a part of the core research area of artificial vision, and an important number of authors have reported methods and applications for people detection and positioning. More recent and therefore less abundant are those works involving the use of modern technologies such as RGB-D sensors, which provide 3D information in form of *depth maps* of scenes. For example, Saputra *et al.* [3] presents an indoor human tracking application using 2 depth-cameras. Although some effort has been made in applications using the above-mentioned technologies, there is a lack of research on the combined use of both types of technologies for positioning purposes.

This article presents a new IPS approach based on the combination of these two different technologies: WPS and *depth maps*, in an active manner. By active combination these authors mean that the developed method puts together and exploits coordinately both types of sensory information: strength of measured wireless signals and *depth maps*.

This approach is particularly advantageous when several users are simultaneously in a room. In this case, the system is able to detect each user with the help of the coordinates of the people located in a *depth map*. WPS approximates the position of the users, but when they are really close, the proposed method is able to deliver a more precise location. This is carried out with the help of user trajectories, which are considered in two ways: WPS trajectory and trajectory of the people in the *depth map*. As demonstrated in the following sections, this combination improves the efficiency of the existing approaches used in WPS.

The article is structured as follows: Section 2 explores existing solutions concerning to positioning, based on WPS, RGB-D sensors, and using both technologies in a joint manner. Section 3 is devoted to describe in detail the basis of the proposed system and how does it works. Section 4 presents the performed experiments and analyzes the obtained results. Finally the Section 5 remarks the main advantages of the presented system and shows future developments based on this method.

2. Overview of related work

Recently, Subbu *et al.* [4] established three types of IPS: *Fingerprinting*, which uses the signals obtained from portable device such as WiFi, sound, light or magnetic fields; *Crowdsensing*, an extension of *fingerprinting* that continuously updates the positioning database; and finally *Dead Reckoning Systems*, using the accelerometer sensor of portable devices to obtain the inertial movement and the magnetometer to obtain the direction of the magnetic field.

WPS is founded on the *fingerprinting* technique [5], in which a map of the environment is created recording various values of *Received Signal Strength Indication* (RSSI) in each point. RSSI is a reference scale used to measure the power level of signals received from a device on a wireless network (usually

WiFi or mobile telephony). This map is used to obtain the position of a user in real-time, comparing the values received from the user's portable device to those stored in the map.

Quan *et al.* [6] show how WPS based on *fingerprint maps* works better than those techniques based on triangulation, like RADAR [7]. This technique [7] records and processes signal strength information at multiple base stations and combines empirical measurements with signal propagation modeling to determine users location by means of the triangulation.

The positioning with *fingerprint map* is carried out in two ways: considering the nearest neighbor, where the Euclidean distances between the live RSSI reading and each reference point fingerprint is calculated for determining the position, and the probabilistic location with Markov, where statistical data of the fingerprint are used to guess the most likely position. The results shown indicate that the nearest neighbor approach works better than the Markov-based one. The triangulation method provides worse results because equations do not transform properly RSSI values into distance, due to the presence of walls and obstacles. Other works have tried to obtain that distance through the use of fuzzy logic [8] or particle filters [9].

Regarding approaches based on fingerprinting, Martin *et al.* [10] study the accuracy of different techniques: *Closest Point*, *Nearest Neighbor in Signal*, *Average Smallest Polygon* and *Nearest Neighbor in Signal and Access Point averages*. Depending on the room, or cell size where the user is situated, the positioning results are different. The successes are between 78% and 87% determining the room where the user is. If the user is in a room and 2x2 meters cells have been created, the successes are between 39% and 48% determining the cell where the user is. When 1x1 meter cells are used, the successes are between 18% and 32%.

Considering the distance between APs and receivers, Kornuta *et al.* [11] analyze the attenuation of the signal produced when the APs are far from the receiver or there are walls or obstacles along the way. Some filters are studied in [12] for attenuating the noise of RSSI. The work [13] studies the combination of WiFi and *Inertial Navigation Systems* (INS) in order to obtain the trajectory of the user. Three sensors are used: gyroscope, accelerometer and an atmospheric pressure sensor. Husen and Lee [14] propose how to obtain the user orientation with a *fingerprint map*.

In the field of people and objects detection, other technologies, such as those based on artificial vision (e.g. RGB-D sensors), have been increasingly used. Ye *et al.* [15] propose to use three Kinect sensors for detecting and identifying several people that are occluded by others in a scene. In [16], authors propose a *smart-cane* for the visually impaired that, with the help of a Kinect sensor, allows to locate objects. The method *Kinect Positioning System* (KPS) is analyzed in [17] aiming to obtain the user position.

These positioning techniques have also been used in Robotics. A noteworthy example can be found in [18], where several *Simultaneous Localization and Mapping* (SLAM) algorithms are proposed for building maps using robots with continuous positioning. Mirowski *et al.* [19] analyze how to generate a *fingerprint map* with an RGB-D sensor mounted on a robot. By means of SLAM, the environment is built recording the measurements RSSI in each point. Also in this field of research, the use of distinct technologies allows to improve the positioning systems. In [20], a robot is located using three different systems: a laser rangefinder, a depth camera, and the RSSI values. Each system is used independently according to the zone where the robot is located.

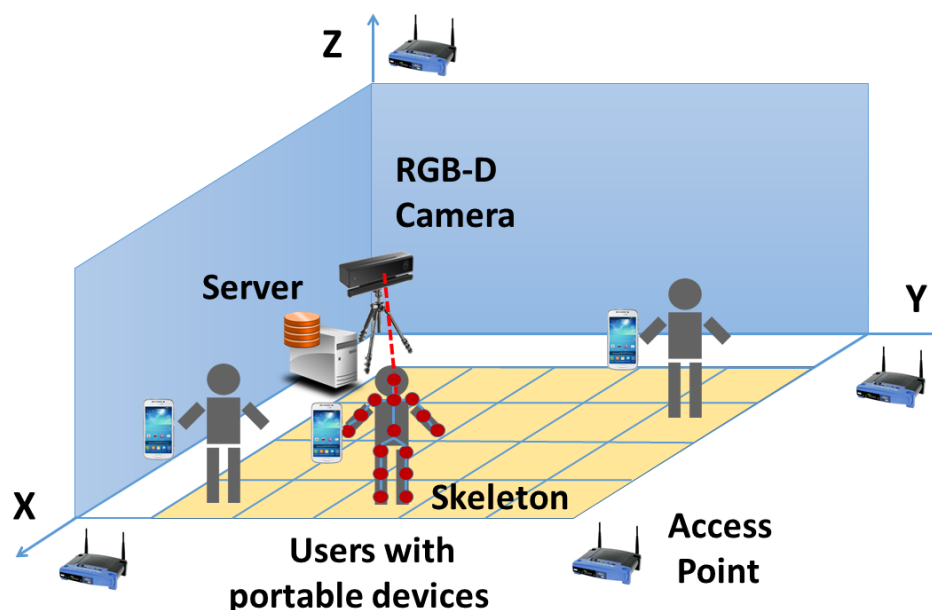
RFID techniques have been proposed for location and tracking of users inside buildings as presented in [21], where authors propose to combine identification and positioning based on RFID with the Kinect sensor for obtaining the precise position of a person inside an environment. In this case, one fix RFID reader is located in the room. Each user carries their own RFID tag while the Kinect sensor obtains the skeletons of two people. Each skeleton is composed of the coordinates of the different joints of a person: neck, shoulders, elbows, knees, etc. Other methods use RFID tags on the floor where the users can know their positions thanks to a RFID reader they carry with them [22].

However, RFID techniques present several disadvantages, such as interferences with materials and devices, and do not provide too precise location results. These inconveniences, among others, have encouraged these authors to find an alternative solution that delivers better results in terms of accuracy.

3. Analysis of the system

The aim of the proposed system is to increase the accuracy of people positioning inside a room. To do that, let us consider a scenario like depicted in Figure 1, which represents the generic framework of the system. One or several persons are assumed to be freely moving around a rectangular working area, carrying their own portable device. Each device receives its corresponding wireless signal from one or more APs strategically located in the working area. One RGB-D sensor is placed in such a way that most of the working area is covered. This device delivers a *depth map* and a color image of the scene, that are used to identify the 3D skeletons of users. The skeletons are obtained by means of the techniques presented in [23] [24], where authors propose new algorithms to quickly and accurately predict 3D positions of body joints from depth images. Those methods form a core component of the Kinect gaming platform. From these skeletons, neck coordinates are extracted aiming to position people in the environment. This part of the body is chosen because it is less prone to be occluded by elements in the scenario. Finally, a server computer is used for controlling the overall process.

Figure 1. Components of the system



3.1. System working description

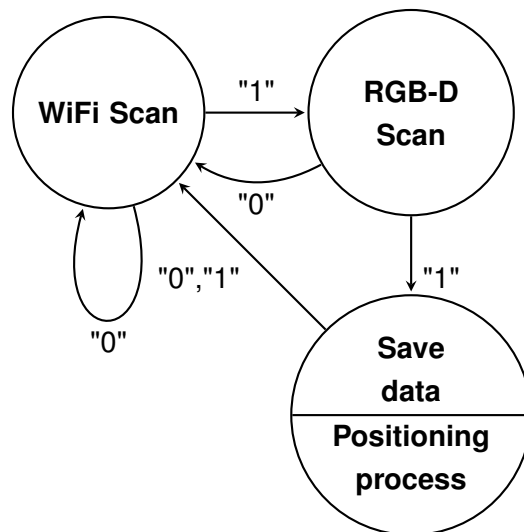
The developed system is divided in two main stages: *Learning* and *Running*.

3.1.1. Learning stage

The main purpose of this stage is to create, for the selected working area, a new database with the processed information coming from the two technologies: WPS and RGB-D. During this stage, the *fingerprint map* associated to one user is created by registering simultaneously the RSSI values obtained by their portable device and the coordinates of their skeleton. The user moves alone around the room in order to match each RSSI scan with each skeleton position. This task is performed in three different steps: *WiFi Scan*, *RGB-D Scan* and *Save data*.

During the *WiFi Scan*, the portable device obtains RSSI values for each AP and sends them to the server. When RSSI values are received, the *RGB-D Scan* is started. This process returns the skeleton of the person detected in the room. The system automatically saves the RSSI data and, additionally, the user coordinates of the skeleton are obtained from the *depth map*. Figure 2 shows the system diagram.

Figure 2. System diagram during Learning/Running stage where values 1 and 0 represent successful or unsuccessful state execution, respectively. "Save data" is the third step during Learning Stage while "Positioning process" is the step during Running Stage



To simplify the positioning process without significant loss of precision, other essential tasks are carried out at the end of this stage: The floor of the working area is divided in rectangular cells and, RSSI data are grouped in each cell, using the cell position of the skeleton.

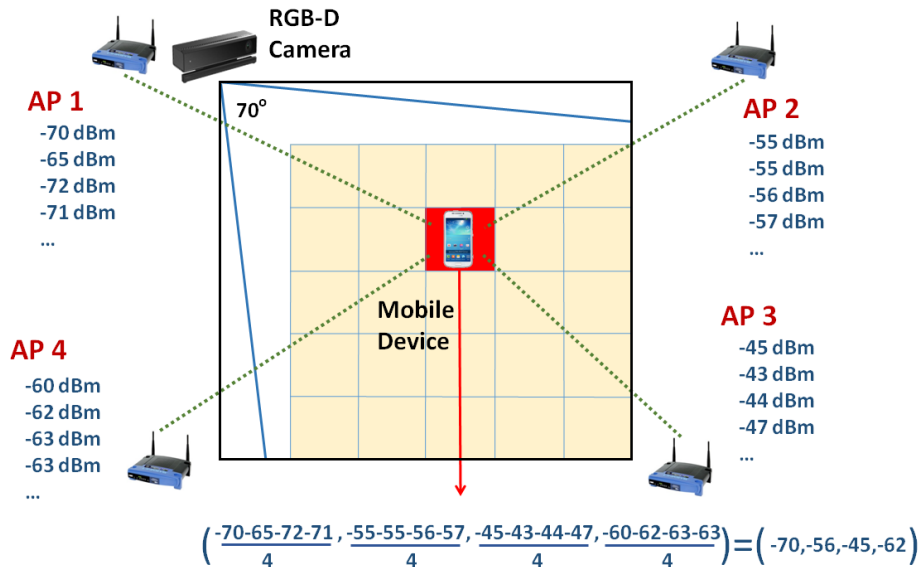
The division in cells is produced when the maximum and minimum coordinates of X and Y (see Figure 1) have been obtained. The coordinates of the skeleton (x, y) deliver the cell where the user is located (c_x, c_y) , according to Equation 1.

$$\begin{cases} c_x = n_x \cdot \frac{x - \min_x}{\max_x - \min_x} \\ c_y = n_y \cdot \frac{y - \min_y}{\max_y - \min_y} \end{cases} \quad (1)$$

where n_x and n_y represent the number of cells in each axis while the variables \max_x , \min_x , \max_y and \min_y represent the highest and lowest values of each axis (obtained from the *depth map*). Note that Z coordinates are not considered as the user position is estimated in 2D.

An RSSI vector is created for each cell, pairing each component to the centroid for all of the RSSI measurements for a certain AP (see Figure 3). This allows reducing RSSI variability.

Figure 3. Obtaining RSSI vector for one cell (example with 4 APs)



3.1.2. Running stage

This stage represents the normal way of working of the system. It is performed by using the three different steps shown in Figure 2 and considers that several users are moving around the room.

While the *WiFi Scan* is running, each user synchronously sends their RSSI values to the web server. When these data are received, the *RGB-D Scan* starts aiming to obtain the skeletons of people detected in the room. Finally, the *Positioning process* estimates the position of each user by combining both data in such a way that each skeleton is linked to each RSSI scan.

In the *Positioning process*, different algorithms are executed depending on the system's running mode, going from the simplest *Basic Mode*, in which only WPS method is applied, to more sophisticated ones, where both type of sensors are combined so that each skeleton is linked to each RSSI scan.

The system stores the different RSSI measurements received from the *WiFi Scan* in a table (see Table 1), and the skeleton coordinates obtained from the *RGB-D Scan* in a different one (Table 2). Note that users A and B in Table 1 are not related to users M and N in Table 2. During the *Positioning process*, the system will be able to decide if A corresponds to M or N and, in the same manner, if B corresponds to M or N. Skeleton coordinates and RSSI data are linked by a time stamp.

Table 1. Example of RSSI Table

Time Stamp	User	BSSID	SSID	RSSI
1 (8:31:57)	A	dc:53:7c:25:2c:36	Router 1	-53
1 (8:31:57)	A	84:9c:a6:fe:0e:34	Router 2	-46
1 (8:31:57)	A	4c:54:99:df:9e:ec	Router 3	-93
1 (8:31:57)	B	dc:53:7c:25:2c:36	Router 1	-74
1 (8:31:57)	B	84:9c:a6:fe:0e:34	Router 2	-40
1 (8:31:57)	B	4c:54:99:df:9e:ec	Router 3	-73
2 (8:31:58)	A	dc:53:7c:25:2c:36	Router 1	-52
2 (8:31:58)	A	84:9c:a6:fe:0e:34	Router 2	-51
2 (8:31:58)	A	4c:54:99:df:9e:ec	Router 3	-89
2 (8:31:58)	B	dc:53:7c:25:2c:36	Router 1	-73
2 (8:31:58)	B	84:9c:a6:fe:0e:34	Router 2	-41
2 (8:31:58)	B	4c:54:99:df:9e:ec	Router 3	-72

Table 2. Example of skeleton coordinates (Neck coordinates).

Time Stamp	User	X	Y	Z
1 (8:31:57)	M	-0,298601	0,035828	1,237208
2 (8:31:58)	N	0,229597	-0,025246	1,738968

The structure of the recorded RSSI data contains the Basic Service Set Identifier (BSSID), the Service Set Identifier (SSID), RSSI and the time stamp. BSSID is formed by the Media Access Control (MAC) of each AP. SSID corresponds to the name used by the APs.

BSSID is used instead of SSID. SSID is informative and can be repeated in WLAN since different APs may have the same network name. RSSI data, SSID and BSSID are collected by the portable devices using the 802.11 layer. At the same time, the portable devices must establish a connection to some accessible network. This can be a WiFi network or a wireless data network of telephony (3G, 4G, etc). The devices send data, via SOAP protocol through the application layer, to a web server. This web server must be connected to the RGB-D camera but it might not be in the same network than the devices since the web services are available on the Internet.

Different RSSI data entries, as well as different skeletons data entries, can be synchronously produced at the same time stamp, as can be observed in both Table 1 and Table 2.

As mentioned before, three different running modes are considered in this work: *Basic Mode*, *Improved Mode without Trajectory* and *Improved Mode with Trajectory*. Their respective features are discussed in next paragraphs.

Basic Mode: WPS only

In this mode, RSSI measurements are obtained from portable devices and compared to the values stored in the *fingerprint* database. During the *Learning stage* the RSSI values of the *fingerprint* were grouped using the centroid of the cells, which reduces RSSI variability.

An error, based on the Euclidean distance between the measured RSSI vector and the RSSI vectors of the centroid of each cell, is calculated. The estimated WPS cell is the one with lowest error. Equation 2 shows how this error is obtained from two RSSI vectors: the first one read by the portable device and the second one corresponding to the centroid of each cell. Each vector has n components corresponding to each AP. $v_{p,new}$ represents the component of the vector for an AP p where the user is located, while v_{p,c_x,c_y} represents the component of the centroid vector for that AP p in each cell (c_x,c_y) .

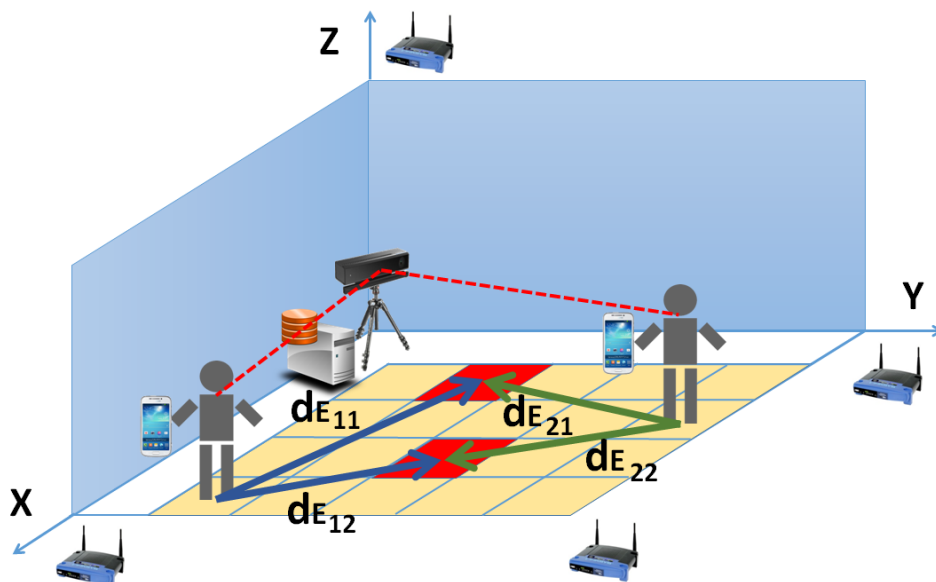
$$e_{c_x,c_y} = \frac{1}{n} \cdot \sqrt{\sum_{p \in \text{AP set}} (v_{p,new} - v_{p,c_x,c_y})^2} \quad (2)$$

Improved Mode without Trajectory: combining WPS with depth maps

In this mode, the information provided by *depth maps* helps determine in which cell the user is located. Furthermore, it is useful for clarifying their exact position. The combination of both methods improves indoor positioning in a simple manner.

Two different cases are studied depending on the number of users inside the room: if there is only one user in the room, the *depth map* allows to obtain the exact position. The portable device provides the right identification of the user.

Figure 4. Scheme of two people in the room



When there are two or more users, as shown in Figure 4, several skeletons are obtained. At the same time, each user sends a group of RSSI measurements to the server. The system initially does not know what skeleton is linked to any particular RSSI data. The proposed method calculates the Euclidean distance between the RSSI data, sent by the smart phones (named WPS cell), and the RSSI centroid of the cell where each skeleton is. Then, the system looks for the best combination between each skeleton

and the RSSI measures obtained from each smart phone. This occurs when the sum of the Euclidean distances reaches the minimum (see Equation 3)

$$\min \sum_{i,j} s_{ij} \cdot d_{E_{ij}} \wedge s_{ij} \in \{0, 1\} \wedge \sum_x s_{xj} = 1 \wedge \sum_y s_{iy} = 1 \quad (3)$$

where s_{ij} represents the links between each skeleton and its WPS cell and can take the values 0 and 1. $d_{E_{ij}}$ represents the Euclidean distance between the skeleton i and the position j , where one user has been detected according to WPS.

Improved Mode with Trajectory: considering the trajectory of the user with WPS and depth maps

In this mode, the combination of *depth maps* and WPS also allows to obtain two different trajectories. The trajectory of the user with WPS represents the cells that the user has previously visited, according to data from WPS. The trajectory of the user in the *depth map* is a group of skeletons obtained for each time stamp. Both trajectories (WPS and skeletons) are synchronized with their time stamps, so when skeletons are received, RSSI values are obtained for all users.

When two or more users are simultaneously in the room, and each one has a different trajectory of WPS and skeleton, the system initially does not know what skeleton is linked to each user. However, it can calculate it according to an extension of Expression 3, as explained in the following.

As mentioned in [25], *Synchronized Euclidean distance* measures the distance between two points at identical time stamps. If two trajectories with different points are obtained at the same time (for each pair), the total error is measured as the sum of the distances between all points (points in WPS trajectory and points in skeleton trajectory) at synchronized time stamps.

Figure 5. Example of trajectories of $n = 2$ users in $m = 4$ time stamps, including Euclidean distances between the combination of $P_i(t)$ and $P'_j(t)$

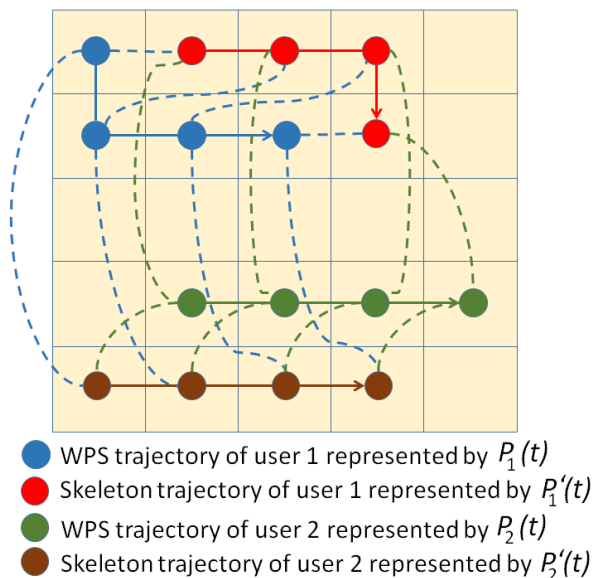


Figure 5 shows the WPS and skeleton trajectories of 2 users at 4 time stamps. $P_i(t)$ represents the WPS position of the user i at the time stamp t . $P'_j(t)$ represents the skeleton position of the user j at the time stamp t . Although Figure 5 represents the trajectory of the user 1 ($P_1(t), P'_1(t)$) and the trajectory of the user 2 ($P_2(t), P'_2(t)$), the trajectories of each user are not linked. The system initially does not know if $P_1(t)$ is associated to $P'_1(t)$ or $P'_2(t)$, and in the same way if $P_2(t)$ is associated to $P'_1(t)$ or $P'_2(t)$. To solve this problem, Expression 4 is used.

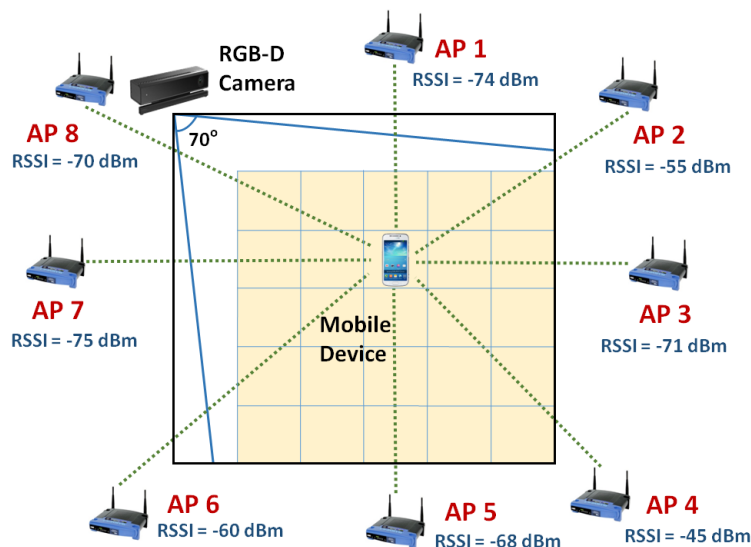
$$\min \sum_{i,j} s_{ij} \cdot \sum_{t=1}^m d_E(P_i(t), P'_j(t)) \wedge s_{ij} \in \{0, 1\} \wedge \sum_x s_{xj} = 1 \wedge \sum_y s_{iy} = 1 \quad (4)$$

This expression 4 takes into account the *Synchronized Euclidean distance* computing the sum of distances between each pair of points (WPS and skeleton trajectory) and looking for the best combination between WPS trajectories and skeleton trajectories to obtain the minimum sum of all distances of all trajectories. Figure 5 shows all of the different *Synchronized Euclidean distances* that are computed for 4 time stamps, where there are two users with two different skeleton trajectories and two different WPS trajectories respectively.

4. Experimentation and results

Several experiments have been developed in a 4.5 x 4.5 meters room where various APs are available (see Figure 6). An RGB-D sensor was located in one of the corners. In this case, eight APs have been located at different positions of the building. Four of them were situated at less than 6 meters from the user. Various users have participated in the process using portable devices, *smartphones* running Android.

Figure 6. Scheme of the room where the tests have been performed



One RGB-D sensor based on time-of-flight technology (ToF), Kinect v2, has been employed in these experiments. This device delivers up to 2 MPx images (1920 x 1080) at 30Hz, and 0.2 MPx *depth maps*

with a resolution of 512x424 pixels. This Kinect camera is connected to a web server where data is saved and processed.

The horizontal field of view of the RGB-D sensor is 70° so, as shown in Figure 6, it is only able to detect people in a section of the room (in yellow). This section has a size of 3.71 x 3.71 meters.

During the *Learning stage*, one user has generated the *fingerprint map* and the matching with the skeletons. The user has moved around the room to produce 1000 different measurements. They have been obtained periodically, sending the RSSI values to a web service hosted on the server. Each time this web service was called, a skeleton scan was performed and the coordinates of the neck were saved, aiming to represent the position of the user.

At the end of the *Learning stage*, the floor has been divided in 25 cells (5 x 5 square cells of 0.74 meters side) and the RSSI centroids for each cell have been calculated. RSSI scans have been grouped according to the distance between their original associated skeleton and the center of each cell.

4.1. Positioning experiments

The results obtained in *Basic Mode* shows that the WPS error of positioning a person inside a room is higher than 2 meters. This high error does not allow to distinguish between different users. For this reason, different experiments have been carried out, with one, two and three users simultaneously to prove the efficiency of the *Improved Mode*.

In the case of one user, the positioning is succeeded in 100% of cases because RGB-D sensor detects just one skeleton. When there are several users simultaneously, RSSI values are synchronously sent to a positioning web service at the same time stamp. The server obtains a skeleton capture of all users present in the room and finally calculates and returns their positions. One result is satisfactory when the system is able to correctly detect the cells where the users are. If each user is situated in a different cell, the system also determines their right positions according to the skeletons.

250 tests have been done, considering 2 or 3 users in the room. The results show that, when trajectory is not used, two users are properly detected in 73% of cases and three users in 46% of cases. Most of the errors are produced when the users are in the same cell. When the trajectories of the users are taken into account, the results improve considerably. As shown in Table 3, a comparative of the results has been done at four different time stamps.

Table 3. Results of different experiments

Use of trajectory	Successes (2 users)	Successes (3 users)
No	73%	46%
Yes	89%	71%

The results show that the best performance is obtained when the users are initially in different cells. When the number of users increases, the performance is lower because there are more skeleton trajectories for each WPS trajectory. But considering that the results are above 71% for three users in a small room (4.5 meters x 4.5 meters), the system delivers an accurate position of the users in most of cases.

5. Conclusions

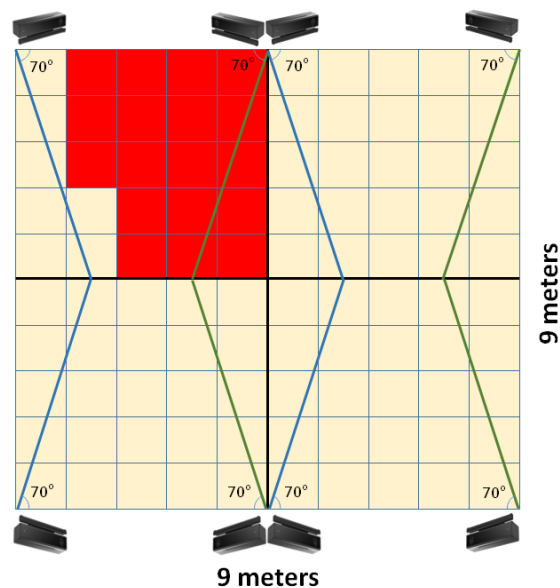
This work presents a new method for indoor positioning based on the combination of WPS with *fingerprinting* and the use of *depth maps*. One RGB-D sensor has been used to obtain the *depth maps* and, subsequently, the skeletons. The combination of both technologies is a simple and economical system that increases the performance of WPS in interiors. The accuracy of WPS detecting users in cells of 2 x 2 meters in a room is lower than 50%. The proposed method allows improving the results until reaching more than 89% for two users and 70% for three.

The combination of WPS and *depth maps* presents some advantages such as: low cost, the use of simple devices (i.e. *smart phones*) and easy installation. Furthermore, the system is not intrusive since the identity of users is not required.

The method proposed is open to use *crowdsensing* [4], because it is possible to add knowledge without doing a new learning. If there is just one user in the environment, the system would be able to recalculate the RSSI centroids for each cell, using the new data obtained from the user (RSSI values and skeleton). This technique would adjust the parameters continuously during system operation.

Besides the number of users, the system is scalable to bigger environments. However, the Kinect sensor has a limited range of a few meters. For this reason, it would be necessary to use more than one device. Figure 7 shows a configuration for a room of 9 meters side. RGB-D sensors are placed aiming to cover the wider angle possible. In this manner, eight cameras would scan the whole room. This figure shows in red the area that would be covered by the top-left sensor. Some sensors cover an overlapped area, which would improve the system accuracy.

Figure 7. System configuration in larger environments



Other non low-cost commercial devices allow obtaining *depth maps* in wider ranges. For example, Peregrine 3D Flash LIDAR Vision System [26] is a lightweight camera able to capture a *depth map* in 5 nanosecond with the help of a Class I laser. It can operate with lenses of 60° and a range over 1 Km.

Despite the fact that this work just estimates the current position of the users, it would be possible to predict their forthcoming path by means of their last trajectories, considering the simultaneous evaluation of WPS and skeleton trajectories.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Acknowledgments

This work has been developed with the help of the research project DPI2013-44776-R of MICINN. It also belongs to the activities carried out within the framework of the research network CAM RoboCity2030 S2013/MIT-2748 of Comunidad de Madrid

References

1. Deak, G.; Curran, K.; Condell, J. A survey of active and passive indoor localisation systems. *Computer Communications* **2012**, *35*, 1939–1954.
2. 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* **2007**, *37*, 1067–1080.
3. Saputra, M.R.U.; Widyawan, W.; Putra, G.D.; Santosa, P.I. Indoor human tracking application using multiple depth-cameras. *Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on. IEEE, 2012*, pp. 307–312.
4. Subbu, K.; Zhang, C.; Luo, J.; Vasilakos, A. Analysis and status quo of smartphone-based indoor localization systems. *Wireless Communications, IEEE* **2014**, *21*, 106–112.
5. Liu, W.; Chen, Y.; Xiong, Y.; Sun, L.; Zhu, H. Optimization of Sampling Cell Size for Fingerprint Positioning. *International Journal of Distributed Sensor Networks* **2014**, *2014*.
6. Quan, M.; Navarro, E.; Peuker, B. Wi-Fi Localization Using RSSI Fingerprinting **2010**.
7. Bahl, P.; Padmanabhan, V.N. RADAR: An in-building RF-based user location and tracking system. *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE. Ieee, 2000, Vol. 2*, pp. 775–784.
8. Feng, X.; Gao, Z.; Yang, M.; Xiong, S. Fuzzy distance measuring based on RSSI in Wireless Sensor Network. *Intelligent System and Knowledge Engineering, 2008. ISKE 2008. 3rd International Conference on. IEEE, 2008, Vol. 1*, pp. 395–400.
9. Svečko, J.; Malajner, M.; Gleich, D. Distance estimation using RSSI and particle filter. *ISA transactions* **2015**, *55*, 275–285.
10. Martin, E.; Vinyals, O.; Friedland, G.; Bajcsy, R. Precise indoor localization using smart phones. *Proceedings of the international conference on Multimedia. ACM, 2010*, pp. 787–790.
11. Kornuta, C.; Acosta, N.; Toloza, J.M. Posicionamiento WIFI con variaciones de fingerprint. *XVIII Congreso Argentino de Ciencias de la Computación, 2013*.
12. Deak, G.; Curran, K.; Condell, J. Filters for RSSI-based measurements in a Device-free Passive Localisation Scenario. *Image Processing & Communications* **2010**, *15*, 23–34.

13. Evennou, F.; Marx, F. Advanced integration of WiFi and inertial navigation systems for indoor mobile positioning. *Eurasip journal on applied signal processing* **2006**, *2006*, 164–164.
14. Husen, M.N.; Lee, S. Indoor human localization with orientation using WiFi fingerprinting. Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication. ACM, 2014, p. 109.
15. Ye, G.; Liu, Y.; Deng, Y.; Hasler, N.; Ji, X.; Dai, Q.; Theobalt, C. Free-viewpoint video of human actors using multiple handheld kinects. *Cybernetics, IEEE Transactions on* **2013**, *43*, 1370–1382.
16. Takizawa, H.; Yamaguchi, S.; Aoyagi, M.; Ezaki, N.; Mizuno, S. Kinect cane: object recognition aids for the visually impaired. Human System Interaction (HSI), 2013 The 6th International Conference on. IEEE, 2013, pp. 473–478.
17. Nakano, Y.; Izutsu, K.; Tajitsu, K.; Kai, K.; Tatsumi, T. Kinect Positioning System (KPS) and its potential applications. International Conference on Indoor Positioning and Indoor Navigation, 2012, Vol. 13, p. 15th.
18. Schindhelm, C.K. Evaluating slam approaches for microsoft kinect. Proc. 2011 The Eighth International Conference on Wireless and Mobile Communications (ICWMC 2012), Venice, 2012, pp. 402–407.
19. Mirowski, P.; Palaniappan, R.; Ho, T.K. Depth camera SLAM on a low-cost WiFi mapping robot. Technologies for Practical Robot Applications (TePRA), 2012 IEEE International Conference on. IEEE, 2012, pp. 1–6.
20. Biswas, J.; Veloso, M. Multi-sensor mobile robot localization for diverse environments. In *RoboCup 2013: Robot World Cup XVII*; Springer, 2014; pp. 468–479.
21. Wang, C.S.; Chen, C.L. RFID-based and Kinect-based indoor positioning system. Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), 2014 4th International Conference on. IEEE, 2014, pp. 1–4.
22. Lee, S.Y.; Min, B.C.; Kim, D.H.; Yoon, J.S.; Kim, D.H. Passive RFID Positioning System Using RF Power Control. In *Robot Intelligence Technology and Applications 2012*; Springer, 2013; pp. 845–853.
23. Shotton, J.; Sharp, T.; Kipman, A.; Fitzgibbon, A.; Finocchio, M.; Blake, A.; Cook, M.; Moore, R. Real-time human pose recognition in parts from single depth images. *Communications of the ACM* **2013**, *56*, 116–124.
24. Barmoutis, A. Tensor body: Real-time reconstruction of the human body and avatar synthesis from RGB-D. *Cybernetics, IEEE Transactions on* **2013**, *43*, 1347–1356.
25. Lawson, C.T.; Ravi, S.; Hwang, J.H. Compression and Mining of GPS Trace Data: New Techniques and Applications. Technical report, Technical Report. Region II University Transportation Research Center, 2011.
26. Peregrine 3D Flash LIDAR Vision System. Advanced Scientific Concepts. <http://www.advancedscientificconcepts.com/products/Peregrine.html> (last access May 12th, 2016).