SWiBluX: Multi-Sensor Deep Learning Fingerprint for Precise Real-Time Indoor Tracking

Alberto Belmonte-Hernández, Gustavo Hernández-Peñaloza, *Member, IEEE*, David Martín Gutiérrez, and Federico Álvarez, *Member, IEEE*

***Abstract*— Indoor/outdoor localization topic has gained a significant research interest due to the wide range of potential applications. Commonly, the Fingerprinting methods for spatial characterization of the environments monitored are employed in deterministic/statistical estimation. However, there are Fingerprint parameters that are generally neglected and can seriously affect the performance yielding to low accurate location. Nowadays, machine and deep learning (DL) methods are employed in this topic due to its ability to approximate complex non-linear models being capable of mitigating the undesirable effects of wireless propagation. In this paper, a complete overview of most influential aspects in Fingerprinting and indoor tracking methods is presented. Furthermore, a novel multi-modal complete tracking system, called SWiBluX, based on statistic and DL techniques is presented. The system relies on relevant feature extraction from available data sources to estimate user's/target indoor position using a multi-phase statistical Fingerprint and DL disruptive approach. In addition, a Gaussian outlier filter is applied to the position estimation model output to further reduce the error in the estimation. The set of experiments performed shows that Fingerprint positioning accuracy estimation can be improved up to 45% resulting in a final estimation error that outperforms related literature.**

***Index Terms*— Indoor positioning, tracking, orientation, fingerprinting, particle filter, wireless, RSSI, IMU, feature vector, machine learning, deep learning, neural network.**

I. Introduction

IN THE recent years, scientific methods for data analysis of human activities have become a relevant topic to extract behavioral patterns [1], [13], [14] which represent valuable metrics for multiple purposes such as research, business, security, health [3] among others [14]. Indoor/outdoor tracking has attracted an extensive focus of research due to the continuous progress in computing capabilities and sensing technologies such as RGB, LIDAR, ultrasound, infra-red, etc.

An extended approach for people tracking is based on Computer Vision techniques using RGB sensors [3]. However, occlusions, lighting changes among others, constrain the estimation accuracy of these methods. Moreover, recent RGB-D and LIDAR sensors (i.e Intel Real Sense or Microsoft Kinect [4]) allow to obtain also depth information attaining a significant accuracy improvement compared to RGB estimation. Nonetheless, short detection range in addition to aforementioned issues decrease the application of these devices. There exist several works that employ ultrasound sensors [5] for tracking. However, lack of accuracy, difficulties associated to multi-object detection and coverage range as well, make these approaches unfeasible for scenarios where individualization is required (i.e places with limited capacity).

Mobile technologies have fostered the development of methods for position estimation and tracking [7] due to two main reasons: *(a)* these devices are widely used among population (smartphones) and *(b)* fulfill most of the aspects concerning privacy.

Advances have empowered personal devices with several wireless technologies (IEEE802.15.1 Bluetooth [30], IEEE802.11g/n WiFi [32] or IEEE802.15.4 [31]) that can be used for indoor position estimation. In particular, two metrics are used for the mentioned technologies: Time-of-Arrival (ToA) and Received Signal Strength Indicator (RSSI) [16]. The former relies on obtaining the time lag of a packet forwarded to an end-point and estimate the distance between devices. However, strong requirements in terms of hardware and synchronization make this proposal hard to implement. The latter approach is most extended as devices are able to obtain the RSSI parameter from packets although propagation effects such as multipath, fading, reflections among others rise a challenge for processing techniques applied [6], [9].

The inclusion of inertial sensors (accelerometers and magnetometers) has increased the set of features that can be applied to prediction models, yielding to meaningful accuracy improvements for person movement monitoring. For the aforementioned techniques (RSSI and ToA), a Wireless Sensor Network (WSN) composed of a set of nodes deployed across the area under monitoring is utilized due to its low-consumption, self-organizing and easy deployment capabilities. Extensive research has been conducted in the literature to address the person tracking problem using RSSI [7], [14], [15], [19]. These proposals are mainly classified into deterministic or statistical methods. On the one hand, deterministic techniques adopt a theoretical model for propagation and geometrically combine the measurements gathered from several nodes to estimate the person position. On the other hand, statistical techniques, such

Manuscript received November 28, 2018; revised January 6, 2019; accepted January 9, 2019. Date of publication January 15, 2019; date of current version April 5, 2019. This work was supported by the European Project: ICT4Life within the H2020 Research and Innovation Program under Grant 690090. The associate editor coordinating the review of this paper and approving it for publication was Prof. Guiyun Tian. *(Corresponding author: Alberto Belmonte-Hernández.)*

The authors are with the Visual Telecommunications Applications Group, Escuela Tecnica Superior de Ingenieros de Telecomunicacion, Universidad Politecnica de Madrid, 28040 Madrid, Spain (e-mail: abh@gatv.ssr.upm.es).

Digital Object Identifier 10.1109/JSEN.2019.2892590

as *Fingerprinting*, characterize the environment by creating a background dataset that contains the measurements at predefined locations. Several Fingerprinting parameters must be adjusted to achieve good performance in position estimation procedure. However, in most of the existing research, there is not a clear description neither of how these parameters were obtained nor how datasets are updated to face with changes in the scenario.

Machine and deep learning algorithms provide novel tools to cope with large datasets, learning complex non-linear patterns from input information [17], [18]. Topics involve data science [20], image classification, detection and recognition [21] among others. Regarding the proposed scenario, Fingerprinting can be treated as a huge dataset of RSSI measurements where these algorithms can be applied [22]–[24]. Once the target position is estimated, tracking is performed by filtering these locations using well known methods such as Kalman Filters (and its nonlinear variants) and Particle Filters [44]–[46] using a reference movement model.

In this paper, a complete framework for indoor estimation and tracking using wireless technologies and inertial sensors is presented. This framework is called SWiBluX from WiFi, Bluetooth and Xbee technologies, it adopts two layers: *(a)* Fingerprinting for position estimation and *(b)* Neural Networks in combination with a novel Gaussian Outliers Filter (GOF) for person tracking. The main SWiBluX contributions are:

- Describe a complete and comprehensive analysis of the most influential parameters in Fingerprint construction.
- Outline a procedure to establish boundaries in parameters selection.
- Draw a Deep Learning (DL) approach for Fingerprint employment in indoor estimation.
- Provide a complete overview of different methodologies in positioning and tracking.
- Present a disruptive and modular architecture for accurate indoor tracking based on a combination of statistical and DL methods.
- Propose a novel Gaussian Outlier Filter for Tracking refinement able to improve the Neural Network performance up to 10%.

The remainder of this paper is organized as follows: section II shows the related work. In section III, the overall indoor tracking problem is formulated. Section IV details the most influential aspects in fingerprinting technique for position estimation. The proposed system is outlined in section V. Results and discussion are described in section VI. Finally, in section VII the main conclusions and future work are drawn.

II. Related Work

The statistical analysis can be further improved based on Fingerprinting methods that overcome the mentioned difficulties, allowing to collect enough data for indoor scenario characterization. Classical propagation models are commonly used, however, harsh and changing indoor conditions make these methods unfeasible due to its low accuracy. Other indoor positioning methods require to adjust several parameters subject to the monitored area, a radio map or additional measurements. Fingerprinting aims to construct a reliable radiomap with a few measurement points, enabling the use more complex techniques to better characterize the final estimated position.

Multiple works have been presented for position estimation using Fingerprinting technique, neglecting the details on how the fingerprinting dataset was calculated [2], [7], [8]. Nonetheless, some information on its implementation was provided in [10], [11], and [47]. In [11], several details regarding the propagation effects have been outlined. However, justification of how the cell number/dimension or device allocation for RSSI collection phase is not provided.

In [12], the effects of user body absorption, antenna orientation and Line Of Sight LOS/NLOS propagation are presented. Results showed how the RSSI attenuation is larger in the presence of human occlusions and it must be considered for Fingerprinting creation. In [13], a four-directions (oriented) Fingerprinting is drawn to enhance the dataset resolution as well as to mitigate the estimation error due to the person/device orientation. Furthermore, an analysis of cell dimension influence showed that large variance of RSSI measures affects the performance of the positioning algorithms. These studies raised challenges on how to reach reliability of the RSSI measurements in Fingerprinting.

Probabilistic methods exploit the RSSI variance for position estimation [34], [40], [41]. Recent approaches based on Machine/Deep Learning techniques are capable of employing the RSSI dataset to learn the optimal model that relates the measurements with the final position estimated. In [37]–[39], methods such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Neural Networks (NNs) are applied. Moreover, works incorporating adaptive updates of Fingerprinting parameters to adjust the models to changes in environment conditions have been depicted in [42] and [43].

Estimated locations are provided as output from the aforementioned techniques. However, these estimations are noisy and neglect temporal dependency. Therefore, using the positions from Fingerprinting stage, several filters can be applied on top to fit the routes obtained to a movement model as well as to reduce the impact from outliers. The most common approaches are Kalman Filters and Particle Filters [44]–[46] which are based on state-space models and combine measurements from different sources to estimate the next state of the system. In fact, Fingerprinting can become an important noise source if the parameters are not properly adjusted.

Probably the most similar works to the one presented here are [23]–[25], where a complete survey of classical, Artificial Neural Networks (ANN) and Reinforcement Learning techniques, attaining a positioning error in the range 1-3 meters. However, SWiBlux extends these approaches by involving additional information extracted from inertial/magnetic sensors to improve the quality of the positioning and tracking algorithms. A custom feature vector is proposed as ANN input to tackle the problems related to Fingerprinting technique. An additional layer to enhance spurious outliers filtering is detailed. In results section VI, it is shown that SWiBluX complete sequential architecture outperforms the literature.

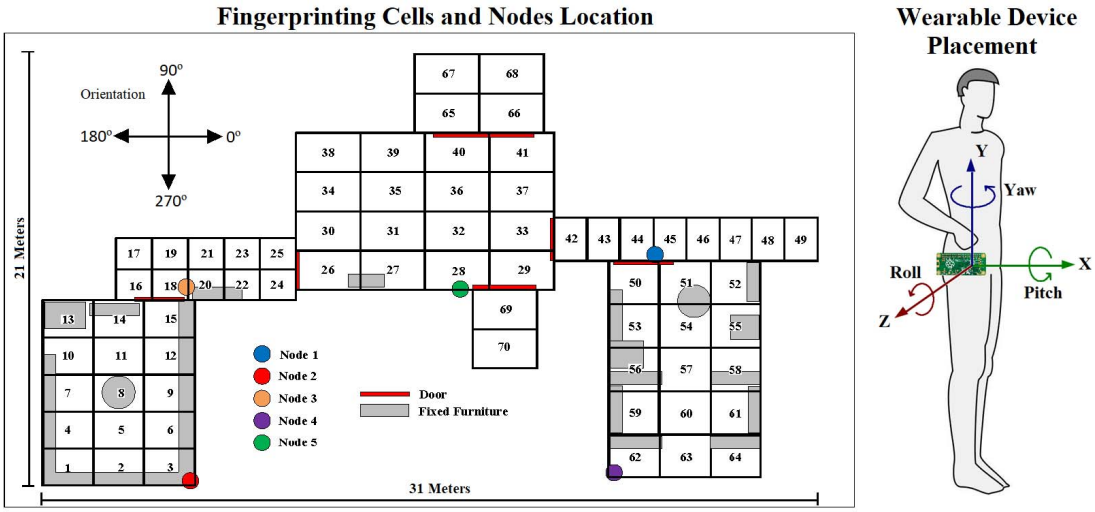

Fig. 1. *Left:* This figure represents the real monitored scenario and the cells distribution for fingerprinting. Five nodes equipped with the selected wireless technologies are deployed across the area (*circles*). Doors are drawn in red and furnitures in gray respectively. *Right:* The reference coordinates system for inertial and magnetic sensors as well as the allocation of devices employed are illustrated.

III. Problem Statement

A typical indoor scenario is characterized in Figure 1 (scenario used in the experiments performed in this work) where the width and depth are denoted by W, D respectively. Additionally, a set of cells $\mathbf{c} = \{\mathbf{p}_l | \ l = 1, \ldots, L\}$ with geometrical center at $\mathbf{p}_l = (x_l, y_l) \in \mathbb{R}^2$ is defined, where c_i is a particular cell. The distance from a cell to another is marked by $d(c_i, c_j)$ using euclidean distance as metric.

It is assumed that a set of nodes are deployed across the indoor scenario under monitoring. These nodes can be modeled as $\mathbf{v} = \{v_{i,\lambda} \mid i = 1, \ldots, I; \lambda = 1, \ldots, \Lambda\}$ and are fully connected. Furthermore, the nodes are located at Cartesian coordinates $\mathbf{p}_i = (x_i, y_i) \in \mathbb{R}^2$. Moreover, the nodes are equipped with sensors from wireless technologies λ able to extract the Received Signal Strength Indicator $RSSI$ from a packet. In particular, sensors from three wireless technologies are considered: IEEE802.11.g/WiFi [32], IEEE802.15.1/Bluetooth v4.0 [30] and IEEE802.15.4/XBee. Let $RSSI_{i,\lambda}$ to describe the RSSI measurements gathered by node i with technology λ.

It is assumed that a/several individuals is/are moving in the scenario wearing mobile devices endowed with at least on of the aforementioned technologies (smartphone, wearable, etc). The main problem addressed in this paper is the individual tracking using the WSN information.

Optimal cell distribution for Fingerprinting is one of the main issues addressed in this work. The most common aspects include: *(a)* identifying the most influential parameters for Fingerprint construction, *(b)* defining some approaches to mitigate the effects of cell distribution, *(c)* data fusion of several wireless technologies, *(d)* analyze and assess deterministic and statistic methods for position estimation based on RSSI dataset as well as Machine/Deep Learning approaches, *(e)* a simple outliers filter based on estimations given by the algorithms and parameters extracted from sensors and *(f)* the implementation of final tracking filtering algorithm.

Raspberry Pi devices were used as nodes, equipped with XBee S1 (protocol IEEE802.15.4, indoor range 30 meters, transmission power 0 dBm) Bluetooth USB Dongle PBU40 (protocol IEEE802.15.1 v4.0 BLE+EDR, indoor range around 10-20 meters, transmission power -3 dBm) and WiFi USB dongle TP-LINK TL-WN722N (protocol IEEE802.11.(b,g,n), indoor range around 40-80 meters, transmission power around 21 dBm), whereas the wearable was endowed with Inertial Movement Unit (IMU) Adafruit 9-DOF LSM9DS0 with accelerometer, gyroscope and magnetometer integrated.

The orientation of the wearable device attached to the body is presented in the right side of Figure 1. The device is allocated on the person hip and axes (X,Y,Z) in the figure match the IMU axes.

IV. Fingerprint Calculation

Aforementioned factors affect the quality of RSSI measurements yielding to large variance in the power level detected by every node, which is critical when the detection is performed in an indoor scenario as the one shown in Figure 1. Fingerprinting is employed to characterize the monitored scenario by splitting it into a set of cells and obtaining RSSI measurements at every cell for a predefined time. However, some important aspects such as the number of cells, the point of interest for cell measurements and the coverage area are generally neglected.

A. RSSI Variance

Initially, it can be argued that the larger the number of cell is, the higher characterization resolution achieved is. However, if the number of cells is large, the statistics could be similar between each other, increasing the estimation errors for Fingerprinting cells with identical behaviors or very close among them. Consequently, it can be also stated that the larger the number of fingerprinting cells is, the higher probability of error in estimation is.

Therefore, an analysis is required in order to select the cell size to prevent the overlap between RSSI measurements

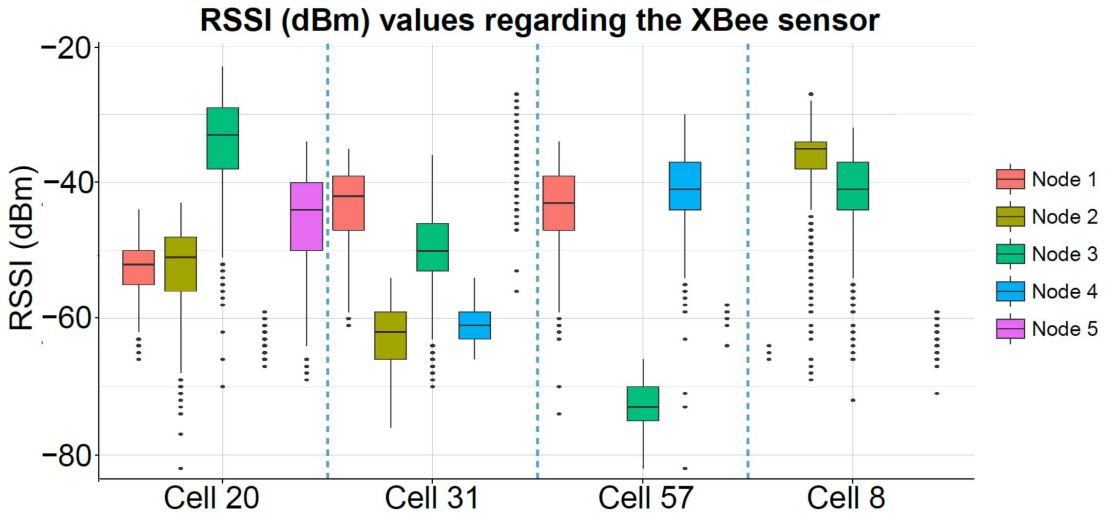

Fig. 2. Example of RSSI variance boxplots in the collected fingerprinting database is presented for each node in different cells selected for a particular wireless technology (XBee). Variance in measurements is large, yielding to significant fluctuations in the signal distribution over a particular location (i.e. a cell in the monitored room).

collected in the nearest cells. In practice, RSSI values overlap for different distances in multiple cells due to the propagation conditions. It can be inferred that the more distances with the same RSSI reading overlapping, the higher error probability in position estimation.

An example of this statement is depicted in Figure 2 where a set of four cells and all the RSSI measurements collected by all nodes are presented. A single node has a coverage range constrained for every particular technology, and any of these covers the whole area. As a consequence, unexpected effects such as a non-uniform data distribution from sensors around this position or even lack of coverage makes difficult to fit a proper propagation pattern for each cell. As an example, for Cell 8, it is shown that measurements collected from Node 2 and 3 (the closest nodes in our deployment) are larger than measurements collected from Node 1 and 5, whereas no measurements from Node 4 (the furthest node with more walls in between).

B. Wireless Technology Signal Attenuation

Wireless technologies used in fingerprinting have different signal distributions across the distance. There exist several approaches for indoor propagation modeling where the signal attenuation is relevant to define the cell size. As an example, cases where signal attenuation in a cell is very low (i.e. almost flat) have a similar distribution giving to unreliable estimations for this area. Based on this statement, the nodes must be allocated accordingly to the area under monitoring avoiding to have these nearly from each other. This effect is mitigated in the presence of multiple wireless technologies as its particular propagation patterns allow to distinguish the locations by fusing the data from each one.

C. Sampling Time Per Cell and Synchronization

The data collection period for every cell is relevant due to the need of gathering enough RSSI measurements to obtain a representative sample. The number of packets forwarded depends on the processing capabilities of transmitting devices. In order to synchronize the measurements collected, it is assumed that the transmitter device is able to embed the timestamp to the measures collected. Fixed length slicing windows of 40 ms were created to collect measurements that are assumed to be synchronized. In the tests performed in this work, the sampling frequency for the considered technologies is up to 25 samples per second and the sampling time per cell was adjusted to 5 minutes.

D. Object or Person Tracking

The absorption of people must be considered for human tracking applications. Several works have addressed this problem trying to model the human absorption [36]. In this work, empirical tests were performed to determine that the average attenuation in detection due to human absorption. In this work, measurements along several locations and different orientations with the device attached to the body and the device fixed without human interactions have been collected. In addition, some measurements with 3-6 persons moving along the monitored area were considered to emulate interference. The mean of the information gathered was used to establish an attenuation equal to 4 dBm.

E. Orientation

The antenna orientation effect has also implications on the statistical RSSI positioning algorithms. In general, such assumption remains only in the symmetric region of the antenna. However, when communications are presented in a non-symmetric region, which is generally the 3D case, there is a considerable variation of RSSI values. Moreover, the radiation pattern is not perfectly symmetric in all directions, having an impact on the signal propagation. An experiment rotating a 802.15.4 wireless device in the same plane, as explained in [35], has confirmed that the orientation of the device is a

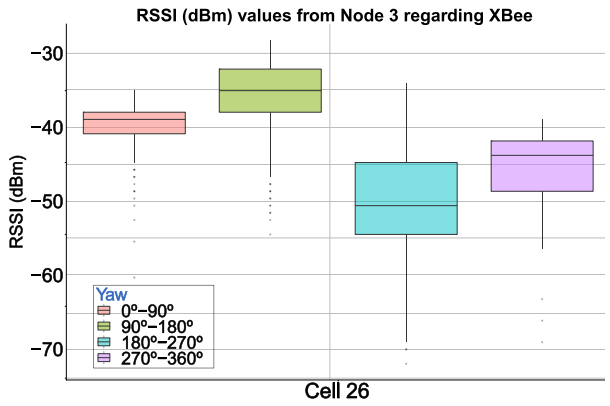

Fig. 3. An example of different orientations in the same cell is presented. The orientations have been concentrated in four angle ranges. The large variability in the measurements depends heavily on the orientation of the device in the cell.

significant attenuation factor in the RSSI value. In Figure 3, it can be observed an example of the measurements collected in a cell rotating over all the angles and representing the RSSI measurements for a single node in four directions. There exists a large variability in the measurements when the device is pointing to the node (−30 dBm) or to the opposite side (−40 dBm).

V. SWIBLUX SYSTEM ARCHITECTURE

Fingerprinting is a method for positioning in particular times, rather than a continuous tracking technique over time. Both, positioning and tracking methods are presented in the architecture drawn in Figure 4. SwiBlux system architecture is composed by several stages: First, a synchronization process over the measurements is needed to create the feature vector. Second, a Deep Learning approach is used to perform the fingerprinting estimation. Third, a Gaussian Filter is proposed to reduce the outliers impact on the estimations. A weighted combination of the estimations after the filter provides an estimated position which is inserted into a Particle Filter to track the person based on a realistic movement model. The use of additional sensors and the proper fusion of diverse sources improves the detection accuracy. The velocity estimation algorithm provides information about how fast a person is moving in the monitored area and the Madgwick [54] filter is used for orientation estimation to improve the final results in some stages of the algorithm.

A. Fingerprinting Position Estimation

In this paper, the literature methods described in section II are compared with Machine/Deep Learning approaches, where the output of this procedure is a probability for all the cells (at center) position based on the feature vector inserted to the algorithm.

$$[RSSI_{1,\lambda}, RSSI_{2,\lambda}, \ldots, RSSI_{I,\lambda}] \tag{1}$$

Data in this format can be analyzed employing different methods such as Distance-based Metrics (i.e. Euclidean,

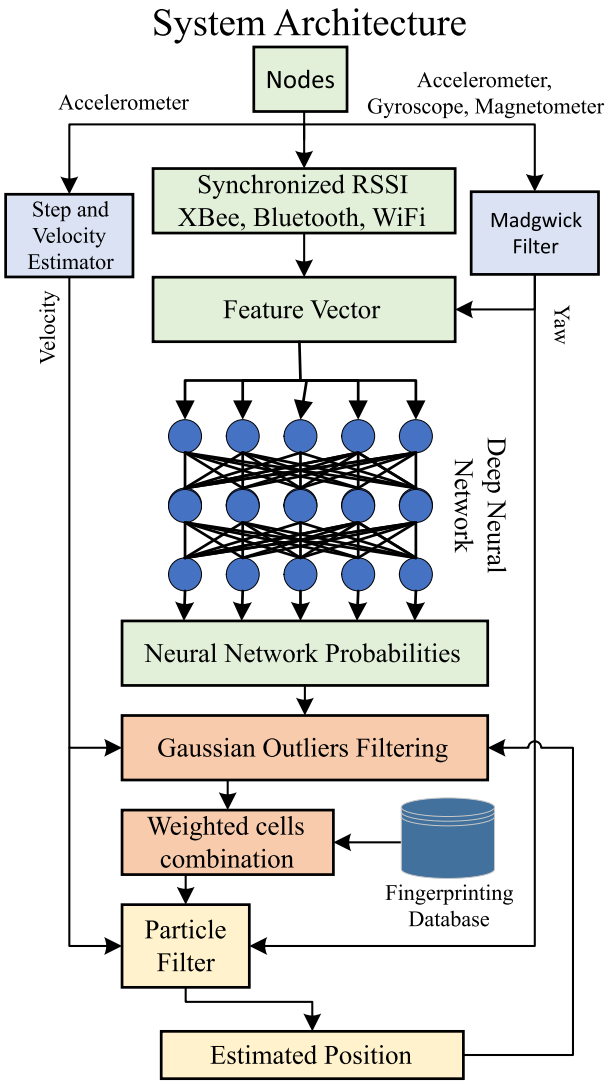

Fig. 4. SwiBlux system architecture is composed by several stages: *First*, a synchronization process an a deep learning approach is used to perform the fingerprinting estimation (green boxes) explained in Section VA. *Second*, a Gaussian filter is proposed to reduce outliers in the estimations (orange boxes) presented in Section VB. *Third*, the step, velocity estimation and absolute orientation (blue boxes) are introduced in Section VC and VD respectively. *Fourth*, particle filter tracks the person based on a realistic movement model giving the final position (yellow boxes) introduced in Section VE.

Mahalanobis, Spearman), K-Nearest Neighbors including its versions (W-KNN) or more sophisticated algorithms such as Kernel Density Estimators. Nonetheless, these methods require some statistical information including both the first and second moments of the distribution as well as the percentiles of RSSI signals in each cell.

Machine/Deep Learning methods can exploit the entire dataset to learn more complex patterns using all the available information. The first step is to gather the RSSI measurements for each cell by every node and create a custom feature vector as shown in the Example 1. The resulting dataset containing these vectors for each cell is used to train the mentioned algorithms.

In Artificial Neural Networks (ANN) training stage, neurons are taught to identify specific incoming signals and whether

activate or not for a particular pattern/output. In related works (see II), the information of the fingerprint in combination with the multiple modalities (from inertial sensors) were employed to train an ANN to learn the most likely cell locations of a target and generalize the probabilities of associated propagation effects. Once the ANN has been trained on samples of the known database, it is able to predict by detecting similar patterns for future input data even when these have not been seen previously. In this work, as input of ANN, the following feature vector (Equation 2) to tackle the Fingerprinting problems in an indoor scenario is presented:

$$[RSSI_{1,X}, RSSI_{2,X}, \ldots, RSSI_{I,X}, RSSI_{1,B}, RSSI_{2,B}, \ldots, RSSI_{I,B}, RSSI_{1,W}, RSSI_{2,W}, \ldots, RSSI_{I,W}, Yaw] \quad (2)$$

where $\{X, B, W\}$ indicates $X = XBee$, $B = Bluetooth$ and $W = WiFi$, I represents the number of nodes deployed and Yaw is the Heading angle estimated by the Madgwick Filter.

This vector includes all available wireless technologies which implies a synchronization stage that allows to employ data from each node deployed for each wireless signal. For instance, in a certain time slot, an XBee packet could have not being received due to the attenuation, whereas the rest of technologies collected properly the RSSI value. Moreover, in spite the Bluetooth signal is weak and the coverage range is not large, the estimation is possible since the rest of technologies have a larger coverage and might be available in cases where Bluetooth is not.

Furthermore, the Yaw/Heading angle introduced in the vector provides relevant information for the orientated RSSI in the cell as shown in Figure 3. The orientation affects the measurements gathered, and these changes are registered during the Fingerprinting creation by a person wearing the device (to take in account body absorption impact) rotating continuously in the cell reference. The corresponding time to collect data on each cell is fixed to 5 minutes. In such period, the continuous rotation during the RSSI gathering allows to cope with multiple possible events to statistically characterize the RSSI cell distribution and the propagation behavior in each direction.

The selected feature vector is used to construct the complete Fingerprinting dataset for ANN training and testing. An ANN with two hidden layers is selected as network structure with a final *Softmax* activation (Figure 6). The activation function selected in the hidden layers is ReLu. By using two hidden layers, it is possible to represent an arbitrary decision boundary with rational activation function that can approximate any smooth mapping to any accuracy. Additional layers can learn more complex representations. Nonetheless, There is a trade-off between architecture layers (Deepness) and the computational complexity of the proposed system. In this work, significant improvements were not attained by increasing the number of layer (>2) whereas the training process was longer and the delay in estimation was perceived.

Figure 5 presents the SWiBluX complete training architecture. Synchronization between the wireless sensors employed is needed and the estimations of Yaw value from inertial and magnetic sensors during database collection step. The proposed scheme can be observed as a classification problem faced with an ANN where RSSI and inertial measures are inserted as input and the cell estimated is provided as output (classes).

Several approaches can be used to select the number of parameters to construct the network [57]. The method to choose both the number of hidden layers and neurons is the *Try and Error Method* that is characterized by repeated, varied attempts which are continued until success or until the agent stops trying. A Forward propagation approach was selected to tune the network parameters as it works better in models with medium/low size of input data as the one presented in this work.

Moreover, weights initialization has played a crucial role in this work due to the fact that *zeros initialization* is heavily penalized during the training process by not learning at all during epochs with more complex optimization algorithms or learning quickly with simpler optimization algorithms. Therefore, *He initialization* [58] was used in this work.

Since the input data is not very sparse, some methods such as *SGD* (Stochastic Gradient Descent), *NAG* (Nesterov Accelerated Gradient) and *Momentum* are preferred. In the complicated task of selecting the optimal learning rate, *SDG* provided the best results on the dataset created in this work and it was chosen consequently.

The learning rate was fixed to a small value of 10^{-3} to find a local or global minima during the optimization by avoiding large jumps over the loss function surface. Further, no overfitting was observed in the experiments, therefore, no regularization method has been included during training.

For the training process, the data was splitted into two subsets, the training and test sets. The ANN model can learn more efficiently with low-range numbers as input. The proposed feature vector was normalized into zero-to-one values due to the fact that all these parameters are always between a range of -92 and -25 for RSSI values and yaw/heading value between $0 - 360$ degrees respectively.

Once the feature is passed, the back-propagation optimizes the weights to minimize the *Categorical Cross Entropy Loss Function* using the selected optimization algorithm.

B. Gaussian Outliers Filtering and Weighted Position Estimation

Machine/Deep Learning based positioning algorithms provide each cell probabilities in the Fingerprinting distribution. In a real scenario, a person moving across the monitored area can stay at different locations besides the cell center. The last layer in the Neural Network usually performs the *Softmax* activation function to obtain a vector with the cell probabilities for a particular input feature vector. However, in case of wrong estimations, a stochastic error is introduced in the next step of the algorithm using only the center of the cell as estimation.

Layer previous to the *Softmax* activation contains the probabilities for each cell of being selected for position estimation. The proposed system removes the *Softmax* activation layer (at testing stage) and applies a Probability Density Estimation

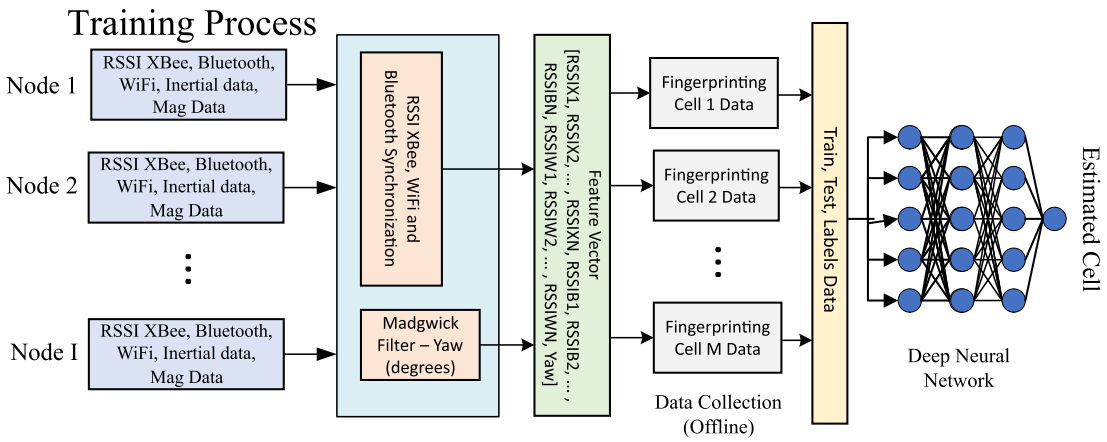

Fig. 5. General scheme of the training process. *From top to bottom*: Fingerprint creation from the N network nodes for each particular technology in addition to the orientation values (Yaw) is depicted. The length of the feature vector inserted into the neural network is $(I(nodes) \times \lambda(technologies)) + 1(Yaw)$.

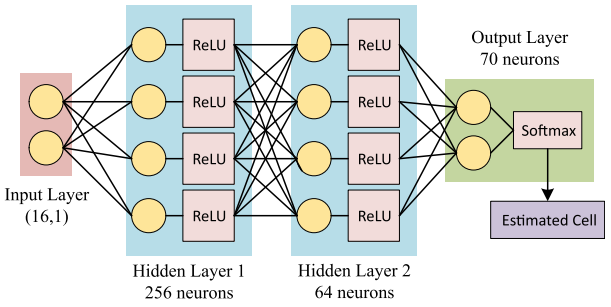

Fig. 6. Deep neural network composed of: an input layer with size the selected feature vector, two hidden layers and the last layer with size equal to the number of cells with a final softmax activation.

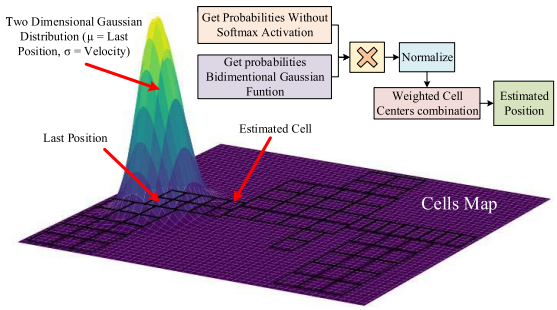

Fig. 7. Example of the Gaussian Outliers Filtering. The bidimensional Gaussian is centered at the previous estimated position and the deviation is the velocity calculated at the current stage. The cell estimated by the neural network, in this case, is wrong so the filter applies a low probability to it. In this manner, the larger probability is assigned to the nearest cells (the ones that the filter covers).

(PDE) filter to separate the unlikely cells (i.e. it is not possible that a person jumps several cells from a previous instant). Additionally, this filter smooths the probabilities of the most likely cells to reach a fine tuning based on the last estimation.

The two-dimensional Gaussian Function (Equation 3) is applied over the outputs of the last layer in the ANN. The key concept is to multiply the outputs of this layer with the ones generated by the two dimensional Gaussian Function. General statistical tools, the *mean* and *deviation* on both spatial axis *(x,y)* are required to generate this function. The mean is obtained using the last position estimated (previous stage) by the Tracking Filter. The deviation is not a constant value and depends on the person movement. Therefore, an adaptive value is chosen to guarantee that the Gaussian Filter width fits with the person movement at each moment.

$$f_X(x_f, y_f) = \frac{1}{(2\pi)|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^\top \Sigma^{-1}(\mathbf{x}-\mu)\right) \tag{3}$$

Person velocity is the key parameter used as the Gaussian Filtering deviation due to its ability to adapt to person changes. To cope with multiple cases, the Gaussian Filter can be wide-spread allowing medium-far cells to be considered into the position estimation when the velocity is high or conversely, to employ narrow filters for slow velocity. This procedure changes the probabilities and provides a normalized vector **p(Gauss)**. Afterwards, a weighted combination of the cells based on these probabilities yield the final estimated location entering to the Tracking Filter as shown in Figure 7.

$$x_g = \sum_{l=1}^{L} \frac{p(Gauss)_i}{\sum p(Gauss)_i} x_l \quad y_g = \sum_{l=1}^{L} \frac{p(Gauss)_i}{\sum p(Gauss)_i} y_l \tag{4}$$

Equation 4 x_g, y_g shows the estimated position before the tracking filter and $p(Gauss)_i$ represents the probability of a cell (with center at location (x_l, y_l)) that can be candidate to the weighted combination.

C. Step Detector and Velocity Estimation

In this work, the accelerometer measurements are used to extract information about the walking/standing state to estimate the velocity based on the signal analysis and patterns. The inertial sensors provide information about the acceleration in the three coordinates axes. An example of measurements gathered during a normal walking behavior are presented in left part of Figure 8. Root Mean Square (RMS) estimation from the three accelerometer axes for the signal analysis was used for processing the raw acceleration data and it is calculated as follows 5:

$$RMS_{acc} = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{5}$$

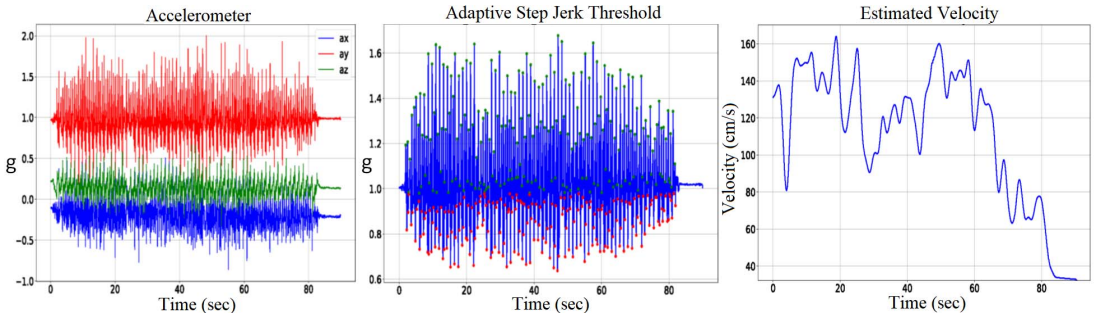

Fig. 8. *Left:* Three axis accelerometer measurements. Regarding the proposed referenced coordinate system it can be shown that the Y axis measurements are around $1g$ representing the gravity force ($9.8\ m/s^2$) in this axis. *Center:* Example of peaks/valleys detection. Red dots represent valleys and green dots represent peaks. It can be observed that the amplitude needs to be large between peaks/valleys to detect them correctly. *Right:* Estimated velocity after Monte Carlo Simulation to get the parameters for the movement model and filtering the output.

To detect steps, maximum (maxima) and minimum (minima) acceleration peaks can be determined by using top and bottom thresholds respectively. However, this method does not work properly due to the large variation of both maxima and minima amplitudes during the walking phases. This issue can cause either lack of peaks or multiple peaks in a single step period. Conversely, the maxima peaks have less amplitude variations than the minimum. To reduce or even eliminate completely this undesirable effect, a peaks/valleys detector with an adaptive threshold is employed. Such algorithm is known as Adaptive Step Jerk Threshold [29] and it is capable to modify the threshold to filter peaks or valleys with no sufficient amplitude or consecutive false peaks/valleys. Using this information, person step length can be determined.

There are empirical equations to obtain the velocity based on these parameters [48]–[50]. However, most of these approaches do not fit properly the velocity model analyzed in this paper. The performance comparison allowed to establish that the best results are obtained using [48]. Equation 6 describes the velocity model based on the output of the peaks/valleys algorithm:

$$L_{step} = K_{conv}\sqrt[4]{A_{max} - A_{min}} \tag{6}$$

where A_{max} is a peak, A_{min} represents the consecutive valleys and K_{conv} is a constant for unit conversion. It is observed that the equation does not adjust this model perfectly due to the fourth square root and the unknown value for the conversion parameter, although this model presents a similar behavior. A Monte Carlo simulation was performed over the root value and the conversion parameter based on different walking measurements of diverse individuals minimizing the error between the real and the estimated velocity.

The last step is to filter the final estimated velocity values. An example of peaks/valey detection along with the estimated velocity by the algorithm over these values are presented in the center and right part of Figure 8.

D. Absolute Orientation Estimation

Common approaches in orientation measurement utilize low cost sensors that lack of accuracy. Nonetheless, the information of accelerometers, gyroscopes and magnetometers can be fused to attain reliable orientation estimations. Three often used choices include the complementary filter [51], the Kalman Filter [52] and the Mahony/Madgwick Filter [53], [54]. In SWiBluX, the Madgwick Filter was chosen due to its good performance when applied over non high precision sensors, low computational complexity and real-time processing.

There exist calibration algorithms to reduce the errors produced by sensors high sensitivity [55]. In this work, the ellipsoid calibration algorithm [56] was selected because of its good trade-off between precision and data required for calibration. This algorithm collects measurements from all the possible sensor positions in the space to correct the bias as well as concentrate the corrected values in a normalized sphere of radius equal to one (Figure 9).

Once the calibration procedure has finished, the absolute orientation can be estimated. This filter uses a quaternion representation, allowing accelerometer and magnetometer data to be used in an analytically derived and optimized gradient-descent algorithm to compute the direction of the gyroscope measurement error as a quaternion derivative.

$$\begin{aligned} Yaw &= \arctan2\left(\frac{2(q_1q_2 + q_0q_3)}{q_0^2 + q_1^2 - q_2^2 - q_3^2}\right) \\ Pitch &= -\arcsin\Big(2(q_1q_3 - q_0q_2)\Big) \\ Roll &= \arctan2\left(\frac{2(q_0q_1 + q_2q_3)}{q_0^2 - q_1^2 - q_2^2 + q_3^2}\right) \end{aligned} \tag{7}$$

The filter output is the estimated quaternion (four dimensional vector) based on the actual sensors measurements. The quaternion vector is employed to calculate the Yaw/Heading, Pitch and Roll as shown in Equation 7.

E. Filtering Tracking With Particle Filter

The tracking algorithm is applied over the estimated position from the previous stage to filter and reduce estimation noise as well as to adjust the positions to an human movement model. It is applied to reduce the fast non-natural movement errors and to fit the model to a like-human movement as shown in Equation 8:

$$\begin{aligned} x_{pf} &= x_p + \hat{v}\cos(\theta)dt \\ y_{pf} &= y_p + \hat{v}\sin(\theta)dt \end{aligned} \tag{8}$$

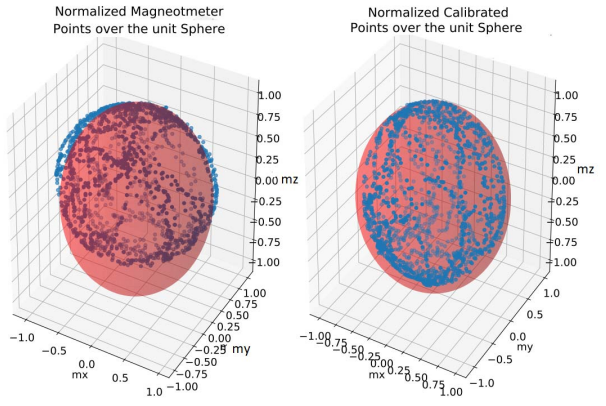

Fig. 9. Magnetometer calibration results. *Left:* uncalibrated normalized data points, *right:* calibrated normalized data points after ellipsoid calibration.

where x_p, y_p are the coordinates of the previous positions estimated, $\hat{v}$ is the estimated velocity and θ is the yaw/heading angle. The employed tracking algorithm is the well-known Particle Filter [59] which is a common choice for applications such as robotics by estimating position, velocity and heading angle in a similar manner to SWiBluX.

The idea behind Particle Filters is that any Probability Density Function (PDF from now on) can be represented as a set of samples (particles). Each particle has a set of values for the state variables representing arbitrary distributions, making it appropriate for non Gaussian, multi-modal PDFs. Therefore, it is feasible to find an approximate representation of a complex model (any arbitrary PDF) rather than an exact representation of a simplified model (Gaussian).

To start the algorithm, an initial belief state, denoted as $p(x_0)$, which is a rough estimation of the PDF is required. For localization problems where no previous knowledge of the distribution is available, it is possible to scatter particles over the whole map. In this work, the particles are scattered around the first estimated position in the first stage with a Gaussian distribution centered at that position with deviation equal to twice the estimated velocity. This is done to cover a wide area of the map in case that the previous estimation might not be reliable. For each time step, the algorithm loop contains three main phases: *prediction*, *update*, and *resampling*.

In the *prediction* step, each particle is displaced according to the motion model. The resulting position from the motion model, will be somewhere in that cloud of particles. The resulting distribution of particles approximates the prior distribution.

During the *update* step, weights are assigned to the sensor measurements according to the probability of observing the sensor measurements from that particles state to each particle. The weights ($\mathbf{w}$) are normalized and its sum is equal to 1 (Equations 9 and 10) to guarantee unbiased decisions.

$$\mathbf{L}_{xy} = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(\frac{-\sqrt{(\mathbf{x}_p - x_g)^2 + (\mathbf{y}_p - y_g)^2}}{2\sigma^2}\right) \quad (9)$$

$$\mathbf{w} = \frac{\mathbf{w} \cdot \mathbf{L}_{xy}}{\sum_{i=1}^{M} w^{(i)} \cdot l_{xy}^{(i)}} \quad (10)$$

where L_{xy} represents the particle importance with respect to the previous estimation and M is the number of particles.

In the *resampling* step, a new set of particles is chosen maintaining the ones with the largest weights. Therefore, unlikely particles (small weights) at the fringe are not chosen whereas the most likely particles (near the center of the cloud) are replicated. As a result, the high-probability region has a high density and properly represents $p(x)$ (the posterior distribution). There are several methods for weights resampling [60], but the Systematic Resample [61] is selected as it faces with the event of particles wandering away from the point and its weight tends (or is equal) to zero. Consequently, a small number of particles implies a reduced number of particles contributing to the approximation of the distribution. Moreover, in order to determine whether resampling is needed or not, the coefficient of variation statistic [62] can be calculated using the following equation 11:

$$CV = \frac{1}{M} \sum_{i=1}^{M} (M w^{(i)} - 1)^2 \quad (11)$$

The effective sample size can be calculated as shown in Equation 12. This parameter describes how many particles have an appreciable weight. Therefore, to check if resampling is necessary, the effective sample size can be tested against the number of particles as follows:

$$ESS = \frac{M}{1 + CV} \quad \text{Resample if ESS} < \tau \cdot M \quad (12)$$

where τ is a threshold between 0 and 1. In this work, this value is selected as 0.5 which means that not resampling is requested in case the half part of the particles exceeds the ESS value.

Finally, the new estimated position is obtained based on the calculated weights and the particles after the prediction step as Equation 13 indicates:

$$x_f = \sum_{i=1}^{M} \mathbf{w} \cdot p_x \quad y_f = \sum_{i=1}^{M} \mathbf{w} \cdot p_y \quad (13)$$

where (p_x, p_y) are the x and y position of each particle respectively and M is the number of particles.

VI. Results on Position Estimation and Tracking

This section shows results on location estimation with the most common methods in Fingerprinting positioning, comparing with the results achievable by different algorithms of both Machine Learning (ML) and Deep Learning (DL) techniques. Furthermore, results on the use of the Gaussian Outliers Filtering and the Weighted Combination of the outputs of this filter after ML /DL methods demonstrates an accuracy improvement in the position estimation. The final location of the person wearing the device is provided by using Particle Filter reducing the final estimated error.

The configuration of the scenario is shown in Figure 1 and represents a typical indoor environment where human activity can be observed and the people flow is controlled. Moreover, harsh propagation conditions arises due to the multi-path fading, the obstacles provoked by the furniture and the wireless

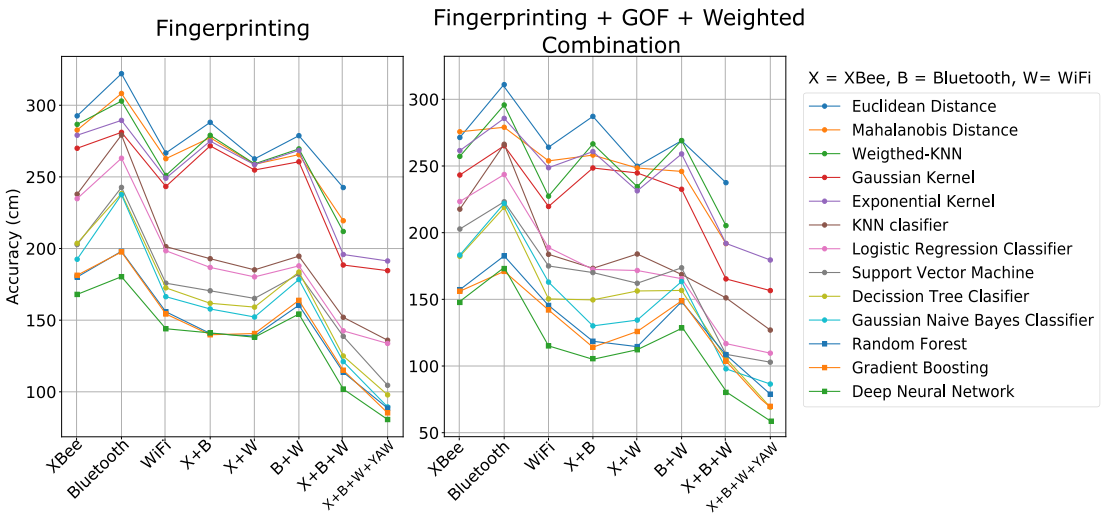

Fig. 10. Estimation accuracy in position estimation using several methods with different feature vectors in two main cases: The first case *(left)* corresponds to perform the normal Fingerprinting estimation based on the selected method, the second *(right)* shows the results after applying the GOF and the weighted combination proposed in this paper. For horizontal axis X, B and W denote XBee, Bluetooth and WiFi technologies respectively.

interferences due to the existence of electronic devices. Further details can be found in [33]. Five nodes $\mathbf{v} = v_1, \ldots, v_5$ are deployed and each node is equipped with three wireless sensors technologies $\Lambda = 3$: IEEE802.15.1 (Bluetooth), IEEE802.11.g (WiFi) and IEEE802.15.4 receiver (XBee).

The monitored area has been splitted into 70 cells of different size varying between 150×150 cm to 240×240 cm. The system is able to collect up to 25 RSSI measurements per second per technology. The collected dataset contains RSSI measurements from each employed wireless technology and Yaw values during 5 minutes in each cell. To consider scenario changes, the measurements have been collected in five different rounds of 1 minute per cell. After the synchronization stage, the database contains 7.5 k feature vectors per cell.

To calculate the accuracy, it is necessary to know where the person is located during a complete route in the monitored area (a.k.a ground-truth). Several routes were collected to get the error in the estimations, comparing the estimated position at each algorithm step and getting the error from the estimated position over the real position using the euclidean distance metric.

Figure 10 shows the accuracy obtained using several methods for fingerprinting estimation taken different approaches for the feature vector in two cases. Accuracy is interpreted as the mean euclidean distance between the estimated position and ground-truth. In the first case, the fingerprinting stage is applied whereas the second case contains the accuracy after applying the proposed Gaussian Outliers Filtering in combination with the Weighted Combination of the probabilities for each cell. These feature vectors contain all the possible combinations of the three technologies employed.

Based on these two figures:

- It can be observed that all methods reach a larger accuracy in the estimations after the application of the GOF followed by the weighted cell combination based on the resulted probabilities.
- Different feature vectors present diverse results. As an example, Bluetooth technology alone has the largest error in all methods as well as it is clear that the combination of technologies outperforms single technology-based estimation. WiFi technology provides a more robust RSSI (with less variance). Thus, all methods work well using this wireless technology in the feature vector.
- The use of the proposed feature vector outperforms the rest of combinations due to the introduction of the Yaw parameter which provides information about the orientation, improving the pattern learning process by Machine/Deep Learning algorithms.
- Deterministic and Probabilistic methods have the worst results compared to learning algorithms. The method including the feature vector proposed in this work can not be applied to distance methods such as Euclidean, Mahalanobis and W-KNN as enough relevant information can not be obtained from the Yaw to be applied.
- The Deep Neural Network achieves the largest accuracy results in both schema. Deep Learning algorithms can learn more complex and non-linear patterns that are limited in the Machine Learning methods.

The velocity model selects the parameters that reach the best fitting to the real measurements with the proposed system. A Monte Carlo simulation over a range of possible values was used to search for the best parameters in several velocity measurements. Finally, the mean of the parameters obtained at each evaluation over real measurements and the final equation are (Figure 11, Equation 14):

$$K_{conv} = 0.81 \quad r_{exp} \approx \frac{1}{3} \quad L_{step} = 0.81(A_{max} - A_{min})^{\frac{1}{3}} \tag{14}$$

These parameters adjust adequately to the proposed method for velocity estimation in different cases when a person is walking slowly, rapidly or even when a person is stopped.

TABLE I

AVERAGE ERROR IN POSITION ESTIMATION (CENTIMETERS) FOR SEVERAL METHODS USING FINGERPRINT AT DIFFERENT STAGES OF THE PROPOSED SYSTEM ARCHITECTURE WITH THE PROPOSED FEATURE VECTOR

Positioning Algorithm \ Proposed System Stage	Fingerprinting	Fingerprinting After Gaussian Outliers Filtering	Fingerprinting + GOF + Weighted Combination	Particle Filter
Deterministic and Probabilistic Common Methods				
Euclidean Distance	259.56	NA	NA	156.34
Mahalanobis Distance	230.56	NA	NA	145.21
Weighted-KNN	210.56	NA	NA	135.67
Gaussian Kernel	184.56	147.78	120.54	102.53
Exponential Kernel	191.32	153.62	132.02	111.36
Machine Learning Methods				
KNN Classifier	135.99	111.21	101.77	87.08
Logistic Regression Classifier	133.76	111.32	99.89	82.32
Support Vector Machine	104.54	90.78	85.67	78.65
Decision Tree Classifier	97.87	90.62	84.98	75.76
Gaussian Naive Bayes Classifier	89.43	82.22	78.45	69.91
Random Forest	88.59	77.79	74.76	62.48
Gradient Boosting	85.32	74.65	71.78	60.56
Deep Learning Methods				
Deep Neural Network	**80.67**	**68.72**	**61.88**	**45.41**

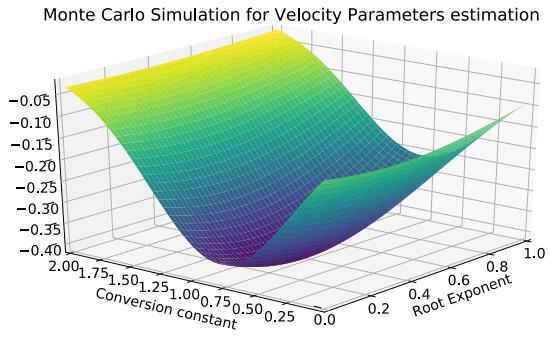

Fig. 11. Monte Carlo simulation over a range of possible values for adjust the the step length model to the detected peaks and valleys by the adaptive step jerk algorithm.

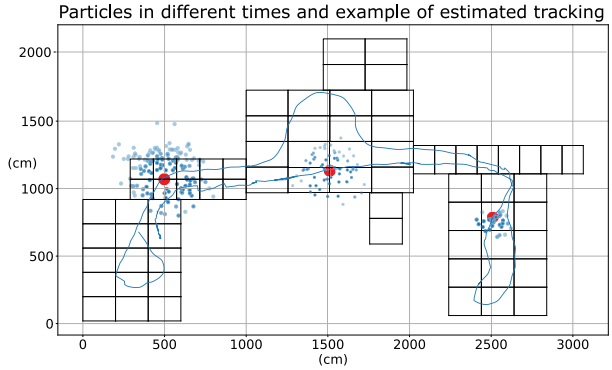

Fig. 12. Particles (blue dots) managed by the Particle Filter at different moments during the tracking. The red dots represent the estimated position at each moment based on the particles whereas blue line depicts the final path estimated using the proposed architecture.

Analyzing the results presented in Table I, it can be observed a significant improvement after each stage in the system architecture:

- GOF improves the final accuracy by filtering critical outliers that can degrade the performance of the final estimations.
- The Weighted Combination based on the resulted probabilities enhances the estimated positon selection, not always placed at the center of the cell, thus reducing the final error.
- Particle Filter tracking algorithm in combination with the estimated parameters (Yaw, velocity) achieves a satisfactory performance with better estimated positions and reducing the average error up to 15 cm.
- Figure 12 represents an example of a person tracked and the particles used by the Particle Filter at three different moments. The number of particles is decreasing over the time according to the selected threshold by the ESS parameter. It can be shown that the inclusion of both Velocity and Yaw provide meaningful information to the filter achieving high accuracy in the estimations.

The ANN has been tested with different configurations looking for high precision while avoiding overfitting. Furthermore, the selection was subject to estimate quickly the position in order to maintain estimations in (near) real time. Four experiments were performed and are presented in Table II. The number of epochs in all configurations is 100. The batch size has been reduced to 64 to learn quickly reducing the overfitting at early stages of the algorithm. Early stopping is used to finish the training if the precision tends to decrease due to overfitting. To ensure a correct dataset splitting in training and validation sets, at learning phase each network was trained ten times with random splitting at each training and the final precision/recall/F1 values are the mean of these ten experiments. The training data size after each splitting is

TABLE II

DIFFERENT NEURAL NETWORKS CONFIGURATION PRESENTED WITH SEVERAL ASSESSMENT PARAMETERS

Eval. Parameters	Precision	Recall	F1 score	Time (ms)
One Hidden Layer 256 neurons				
Neural Network 1	78%	80%	79%	0.896
Two Hidden Layers 256, 64 neurons				
Neural Network 2	89%	88%	87%	1.235
Two Hidden Layers 512, 256 neurons				
Neural Network 3	90%	81%	85%	2.146
Three Hidden Layers 512, 256, 64 neurons				
Neural Network 4	93%	72%	83%	3.689

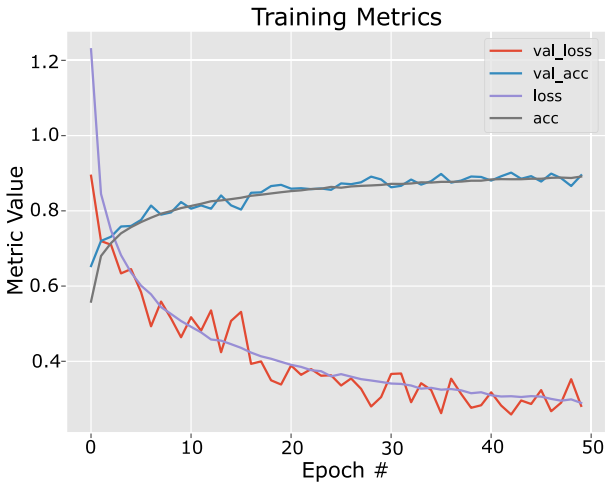

Fig. 13. Training loss and accuracy for training and validation sets in two hidden layers (256, 64 neurons respectively) neural network configuration.

the 85% of the total dataset collected and the test/validation data has the 15% of the total measurements.

- Configuration four (in table II) reaches the best precision but the recall is not large. This is a symptom of the starting of overfitting since the number of neurons in the network layers is large regarding the low size of the input feature vector. It means that the ANN learns quickly the patterns and after some epochs, it *overfits* the final estimations reducing the performance.
- The first configuration presented a short processing time and an acceptable precision, although it is lower than precision obtained by configurations two and three.
- Configurations two and three attain the best precision and recall results. However, configuration two has been selected for the experiments due to its reduced the computation time.

Configuration two has been selected for the experiments regarding the precision of the estimations and the computation time (a simple schema of this network is presented in Figure 6). Both precision and loss during training are depicted in Figure 13. The loss value decreases after each iteration in the training set and the loss in the validation set follows the same decreasing pattern. The precision grows up in training as depicted in the figure but early stopping ended the training process at epoch 49 as the next iteration precision and loss do not decrease.

The entire dataset containing fingerprinting measurements for every technology, as well as example routes are totally available for testing purposes and comparison with the work described here. The scripts for parsing the data are provided and the guidelines in the README file. This information can be downloaded from *http://www.gatv.ssr.upm.es/~abh/*

VII. CONCLUSIONS AND FUTURE WORK

A complete overview on Fingerprint construction and assessment has been presented in this paper. On the one hand, this paper showed the impact of different positioning methods using RSSI Fingerprinting technique on the estimation accuracy. On the other hand, a custom filter was proposed to improve the position estimation results and finally, a Tracking Algorithm was applied to adjust the measurements to a known movement model.

The Fingerprint dataset was constructed taken into account the Fingerprint parameters for diverse situations. Important aspects such as the orientation and the human presence were considered and it has been shown that a significant improvement is attained compared to other estimation procedures. These parameters allow to obtain a consistent feature vector for training Machine/Deep Learning algorithms. Furthermore, the inclusion of the Yaw/Heading information in the feature vector enhances ML/DL training accuracy up to 15%. Final results in Fingerprinting estimated position using learning algorithms outperform the common deterministic and probabilistic methods up to $45-50\%$.

The proposed Gaussian Outliers Filter with the weighted combination of the output probabilities yields to an increase the accuracy around $10-15\%$ in Fingerprinting position estimation by avoiding impossible states or low likely estimations.

Deep Learning approaches provide better results than Machine Learning algorithms when the input data is well conditioned and representative as shown in section VI.

Particle Filter is applied to adjust the final estimations to a realistic movement model. The use of the estimated velocity and the Yaw/Heading angles outperforms the final accuracy presented by all the methods without applying a Tracking Algorithm. An improvement of 25% is achieved after using this filter over the Fingerprinting estimations.

As future work, a study of the impact of varying the number of nodes in the network can be performed. Moreover, a deep analysis of techniques for Fingerprint autorecalibration of over time can yield fast and adaptive systems that respond to the indoor scenario changes. Novel techniques such as Reinforcement Learning can be applied to generate new samples using the previous collected data. Additionally, the research focus on the data fusion of multiple modalities from WSN with visual and depth sensors is relevant to overcome the problems associated to these sensors. An appropriate fusion can solve several tracking problems with cameras such as occlusions, lighting conditions and coverage among others.

REFERENCES

[1] J. Manyika *et al.*, *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. New York, NY, USA: McKinsey Global Institute, 2011.

[2] S. Yiu, M. Dashti, H. Claussen, and F. Perez-Cruz, "Locating user equipments and access points using RSSI fingerprints: A Gaussian process approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[3] G. Tao, P. S. Archambault, and M. F. Levin, "Evaluation of Kinect skeletal tracking in a virtual reality rehabilitation system for upper limb hemiparesis," in *Proc. Int. Conf. Virtual Rehabil. (ICVR)*, Philadelphia, PA, USA, Aug. 2013, pp. 164–165.

[4] W. Zeng, "Multimedia at work," Univ. Missouri, Microsoft Res., St. Louis, Missouri, Tech. Rep., 2012.

[5] S. Sosa-Sesma and A. Perez-Navarro, "Fusion system based on WiFi and ultrasounds for in-home positioning systems: The UTOPIA experiment," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2016, pp. 1–8.

[6] X. Tian, R. Shen, D. Liu, Y. Wen, and X. Wang, "Performance analysis of RSS fingerprinting based indoor localization," *IEEE Trans. Mobile Comput.*, vol. 16, no. 10, pp. 2847–2861, Oct. 2017.

[7] J. Y. Zhu, A. X. Zheng, J. Xu, and V. O. K. Li, "Spatio-temporal (S-T) similarity model for constructing WIFI-based RSSI fingerprinting map for indoor localization," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Busan, South Korea, Oct. 2014, pp. 678–684.

[8] M. Quan, E. Navarro, and B. Peuker, "Wi-Fi localization using RSSI fingerprinting," California Polytech. State Univ., San Luis Obispo, CA, USA, Tech. Rep., 2010.

[9] X. Tian *et al.*, "Improve accuracy of fingerprinting localization with temporal correlation of the RSS," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 113–126, Jan. 2018.

[10] Y. Xu, S. L. Lau, R. Kusber, and K. David, "An experimental investigation of indoor localization by unsupervised Wi-Fi signal clustering," in *Proc. Future Netw. Mobile Summit (FutureNetw)*, Berlin, Germany, Jul. 2012, pp. 1–10.

[11] Y.-H. Li, W.-S. Tang, S.-C. Wang, and P. Hui, "A new method for yielding a database of hybrid location fingerprints," in *Proc. 12th Int. Conf. Intell. Comput. (ICIC)*, Lanzhou, China, Aug. 2016, pp. 212–221.

[12] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *Proc. 1st Annu. Int. Conf. Mobile Ubiquitous Syst., Netw. Services (MOBIQUITOUS)*, Boston, MA, USA, 2004, pp. 14–23.

[13] M. N. Husen and S. Lee, "Indoor Human localization with orientation using WiFi fingerprinting," in *Proc. 8th Int. Conf. Ubiquitous Inf. Manage. Commun. (ICUIMC)*, 2014, Art. no. 109.

[14] A. Farshad, J. Li, M. K. Marina, and F. J. Garcia, "A microscopic look at WiFi fingerprinting for indoor mobile phone localization in diverse environments," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2013, pp. 1–10.

[15] C. Luo, H. Hong, M. C. Chan, J. Li, X. Zhang, and Z. Ming, "MPiLoc: Self-calibrating multi-floor indoor localization exploiting participatory sensing," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 141–154, Jan. 2018.

[16] Q. Dong and W. Dargie, "Evaluation of the reliability of RSSI for indoor localization," in *Proc. Int. Conf. Wireless Commun. Underground Confined Areas*, Clermont Ferrand, France, Aug. 2012, pp. 1–6.

[17] R. Choudhary and H. K. Gianey, "Comprehensive review on supervised machine learning algorithms," in *Proc. Int. Conf. Mach. Learn. Data Sci. (MLDS)*, Noida, India, Dec. 2017, pp. 37–43.

[18] Soniya, S. Paul, and L. Singh, "A review on advances in deep learning," in *Proc. IEEE Workshop Comput. Intell., Theories, Appl. Future Directions (WCI)*, Kanpur, India, Dec. 2015, pp. 1–6.

[19] L. Molina, T. Kerdoncuff, D. Shehadeh, N. Montavont, and A. Blanc, "WMSP: Bringing the wisdom of the crowd to WiFi networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 12, pp. 3580–3591, Dec. 2017.

[20] A. L'eureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine learning with big data: Challenges and approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017.

[21] M. Pak and S. Kim, "A review of deep learning in image recognition," in *Proc. 4th Int. Conf. Comput. Appl. Inf. Process. Technol. (CAIPT)*, Kuta Bali, Indonesia, Aug. 2017, pp. 1–3.

[22] X. Wang, L. Gao, and S. Mao, "CSI phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 1113–1123, Dec. 2016.

[23] X. Gan, B. Yu, L. Huang, and Y. Li, "Deep learning for weights training and indoor positioning using multi-sensor fingerprint," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sapporo, Japan, Sep. 2017, pp. 1–7.

[24] D.-I. Nastac, F. A. Iftimie, O. Arsene, V. Ilian, and B. Cramariuc, "Indoor positioning WLAN based fingerprinting as supervised machine learning problem," in *Proc. IEEE 23rd Int. Symp. Design Technol. Electron. Packag. (SIITME)*, Constanta, Romania, Oct. 2017, pp. 194–199.

[25] L. Yu, M. Laaraiedh, S. Avrillon, and B. Uguen, "Fingerprinting localization based on neural networks and ultra-wideband signals," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Bilbao, Spain, Dec. 2011, pp. 184–189.

[26] A. Estrada, D. Efimov, and W. Perruquetti, "Position and velocity estimation through acceleration measurements," in *Proc. 19th IFAC World Congr.*, Cape Town, South Africa, Aug. 2014.

[27] L. Wang, Y. Sun, Q. Li, and T. Liu, "Estimation of step length and gait asymmetry using wearable inertial sensors," *IEEE Sensors J.*, vol. 18, no. 9, pp. 3844–3851, May 2018.

[28] Y. Kato, H. Nagano, M. Konyo, and S. Tadokoro, "Wearable gait logging system attached on ankles to estimate foot steps and trajectories," in *Proc. IEEE/SICE Int. Symp. Syst. Integr.*, Taipei, Taiwan, Dec. 2017, pp. 294–299.

[29] D. Murray and R. Bonick. *Adaptiv: An Adaptive Jerk Pace Buffer Step Detection Algorithm*. Accessed: Jul. 7, 2018. [Online]. Available: https://github.com/danielmurray/adaptiv

[30] *Bluetooth Technology*. Accessed: Sep. 22, 2018. [Online]. Available: www.bluetooth.com/develop-with-bluetooth

[31] *IEEE802.15.4 WPAN Task Group 4*. Accessed: Sep. 22, 2018. [Online]. Available: http://www.ieee802.org/15/pub/TG4.html

[32] *IEEE 802.11 Wireless Local Area Networks*. Accessed: Sep. 22, 2018. [Online]. Available: http://www.ieee802.org/11/

[33] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[34] S. He, S.-H. G. Chan, L. Yu, and N. Liu, "Maxlifd: Joint maximum likelihood localization fusing fingerprints and mutual distances," *IEEE Trans. Mobile Comput.*, to be published, doi: 10.1109/TMC.2018.2841842.

[35] M. Singh and P. M. Khilar, "Actuating sensor for determining the direction of arrival using maximal RSSI," *Int. J. Sci. Technol. Res.*, vol. 3, no. 8, pp. 182–189, Aug. 2014.

[36] M. Klemm and G. Troester, "EM energy absorption in the human body tissues due to UWB antennas," *Prog. Electromagn. Res.*, vol. 62, pp. 261–280, 2006, doi: 10.2528/PIER06040601.

[37] H. Ahmadi and R. Bouallegue, "Exploiting machine learning strategies and RSSI for localization in wireless sensor networks: A survey," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Valencia, Spain, Jun. 2017, pp. 1150–1154.

[38] J. Torres-Sospedra *et al.*, "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems," in *Proc. 5th Int. Conf. Indoor Positioning Indoor Navigat.*, Oct. 2014, pp. 261–270.

[39] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, Mar. 2015, pp. 1666–1671.

[40] Z. Li, D. B. Acuña, Z. Zhao, J. L. Carrera, and T. Braun, "Fine-grained indoor tracking by fusing inertial sensor and physical layer information in WLANs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–7.

[41] H. Yang *et al.*, "Smartphone-based indoor localization system using inertial sensor and acoustic transmitter/receiver," *IEEE Sensors J.*, vol. 16, no. 22, pp. 8051–8061, Nov. 2016.

[42] S. Cai, W. Liao, C. Luo, M. Li, X. Huang, and P. Li, "CRIL: An efficient online adaptive indoor localization system," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4148–4160, May 2017.

[43] X. He, J. Li, and D. Aloi, "WiFi based indoor localization with adaptive motion model using smartphone motion sensors," in *Proc. Int. Conf. Connected Vehicles Expo (ICCVE)*, Vienna, Austria, Nov. 2014, pp. 786–791.

[44] B. J. Mohd, I. Amro, and A. Alhasani, "Indoor Wi-Fi tracking system using fingerprinting and Kalman filter," in *Proc. IEEE Jordan Conf. Appl. Elect. Eng. Comput. Technol. (AEECT)*, Aqaba, Jordan, Oct. 2017, pp. 1–6.

[45] L. Khalil and P. Jung, "Scaled unscented Kalman filter for RSSI-based indoor positioning and tracking," in *Proc. 9th Int. Conf. Next Gener. Mobile Appl. Services Technol.*, Cambridge, U.K., Sep. 2015, pp. 132–137.
[46] N. Li, J. Chen, Y. Yuan, and C. Song, "A fast indoor tracking algorithm based on particle filter and improved fingerprinting," in *Proc. 35th Chinese Control Conf. (CCC)*, Chengdu, China, Aug. 2016, pp. 5468–5472.
[47] X. Tian, W. Li, Y. Yang, Z. Zhang, and X. Wang, "Optimization of fingerprints reporting strategy for WLAN indoor localization," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 390–403, Feb. 2018.
[48] H. Weinberg, "Using the ADXL202 in pedometer and personal navigation applications," Analog Devices, Norwood, MA, USA, Tech. Rep., 2002.
[49] Q. Tian, Z. Salcic, K. I.-K. Wang, and Y. Pan, "A multi-mode dead reckoning system for pedestrian tracking using smartphones," *IEEE Sensors J.*, vol. 16, no. 7, pp. 2079–2093, Apr. 2016.
[50] I. Bylemans, M. Weyn, and M. Klepal, "Mobile phone-based displacement estimation for opportunistic localisation systems," in *Proc. 3rd Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol. (UBICOMM)*, Sliema, Malta, Oct. 2009, pp. 113–118.
[51] W. T. Higgins, "A comparison of complementary and Kalman filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-11, no. 3, pp. 321–325, May 1975.
[52] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.
[53] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, Jun. 2008.
[54] S. Madgwick, "An efficient orientation filter for inertial and inertial/magnetic sensor arrays," Dept. Mech. Eng., Univ. Bristol, Bristol, U.K., Tech. Rep., 2010, vol. 25.
[55] Y. Wu and W. Shi, "On calibration of three-axis magnetometer," *IEEE Sensors J.*, vol. 15, no. 11, pp. 6424–6431, Nov. 2015.
[56] S. A. H. Tabatabaei, A. Gluhak, and R. Tafazolli, "A fast calibration method for triaxial magnetometers," *IEEE Trans. Instrum. Meas.*, vol. 62, no. 11, pp. 2929–2937, Nov. 2013
[57] G. Panchal and M. Panchal, "Review on methods of selecting number of hidden nodes in artificial neural network," *Int. J. Comput. Sci. Mobile Comput.*, vol. 3, no. 11, pp. 455–464, 2014.
[58] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1026–1034.
[59] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[60] L. Tiancheng, B. Miodrag, and D. M. Petar, "Resampling methods for particle filtering: Classification, implementation, and strategies," *Signal Process. Mag.*, vol. 32, no. 3, pp. 70–86, May 2015.
[61] B. G. Sileshi, C. Ferrer, and J. Oliver, "Particle filters and resampling techniques: Importance in computational complexity analysis," in *Proc. Conf. Design Archit. Signal Image Process.*, Cagliari, Italy, Oct. 2013, pp. 319–325.
[62] L. Martino, V. Elvira, and F. Louzada, "Alternative effective sample size measures for importance sampling," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Jun. 2016, pp. 1–5.

Alberto Belmonte-Hernández received the degree in telecommunication engineering and the master's degree in communication systems from the Universidad Politécnica de Madrid (UPM) in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree. He is currently with the Visual Telecommunications Applications Group, UPM. He is actively working on machine/deep learning algorithms applied to sensors and image for pattern detection and recognition. He has been developing technical parts in national and EU projects.

Gustavo Hernández-Peñaloza (M'07) received the degree in telecom engineering from the Universidad Santo Tomás (Col) in 2007 and the M.Sc. degree in telecommunication technologies, system and networks from the Universidad Politécnica de Valencia in 2009. He is currently pursuing the Ph.D. degree with the Universidad Politécnica de Madrid (UPM). From 2010 to 2013, he was an Associate Research Fellow with the Universidad de Valencia. He is currently with the Visual Telecommunications Applications Group, UPM. He has been participating with different technical developments in several national and EU projects.

David Martín Gutiérrez received the bachelor's degree in audio-visual system engineering from Carlos III University in 2016. He is currently pursuing the master's degree in signal processing and machine learning for big data with Politécnica University. He was with Everis S.L, Spain, as a Software Developer. After, he started to work with Ixion Industry and Aerospace S.L, as a Data Sensor Fusion Engineer. He is with GATV as a Pre-Doctoral Researcher, where he develops artificial intelligent algorithms for different EU projects.

Federico Álvarez (M'07) received the degree (Hons.) in telecom engineering and the Ph.D. *(cum laude)* degree from the Universidad Politécnica de Madrid (UPM) in 2003 and 2009, respectively. He is currently an Assistant Professor and a member of the Visual Telecommunications Applications Group. He had participated in national and international standardization groups. He co-authored 20 papers in refereed journals, more than 10 books and book chapters, and more than 40 papers in international conferences. His current interests include multimedia processing and analytics, networked media, and multi-sensor fusion.