Journal of Ambient Intelligence and Smart Environments 7 (2015) 563–578
DOI 10.3233/AIS-150326
IOS Press

A hybrid radio/accelerometric approach to arm posture recognition

Matteo Giuberti [a,*,**], Marco Martalò [b,c] and Gianluigi Ferrari [c]
[a] *Xsens Technologies B.V., Pantheon 6a, 7521 PR, Enschede, The Netherlands*
E-mail: matteo.giuberti@xsens.com
[b] *E-Campus University, Novedrate (CO), Italy*
[c] *Department of Information Engineering, University of Parma, Parco Area delle Scienze, I-43123, Parma, Italy*
E-mails: marco.martalo@unipr.it, gianluigi.ferrari@unipr.it

Abstract. In this paper, we investigate the feasibility of a hybrid radio/accelerometric approach to perform arm posture recognition. A radio fingerprinting-based approach, through measurements of the Received radio Signal Strengths (RSSs) from anchor nodes, is first used to localize the positions (among a set determined during a training phase) of target nodes properly placed on a user arm. Accelerometric signals generated by the target nodes are then used to estimate the pitch of every device in order to refine the radio fingerprinting results and perform posture recognition, i.e., "continuous" estimation of the positions of the target nodes. We experimentally investigate, through a SunSPOT wireless sensor network testbed, different fingerprinting-based localization algorithms, namely deterministic and probabilistic. In each case, the system parameters are optimized by minimizing a properly defined Position Error (PE). Finally, a comparison between the performance of the proposed system and that of a low-cost optical arm posture recognition system (namely, Kinect) is presented.

Keywords: Radio fingerprinting, accelerometers, arm posture recognition, experimental testbed, Kinect

1. Introduction

Localization systems are becoming more and more important in pervasive wireless technologies for their roles in location-aware services. The Global Positioning System (GPS) has been one of the milestones for outdoor localization [1]. However, its use for indoor localization is often impaired by phenomena typical of indoor scenarios, such as reflections, multipath, and fading. Therefore, indoor localization systems, which do not rely on the use of GPS, have been developed. In particular, recent research has been devoted to the so-called "fingerprinting" technique [2–4]. Fingerprinting is a localization technique where target nodes' positions are estimated on the basis of measurements, by reference nodes (anchor nodes), of the Received radio Signal Strengths (RSSs) from target nodes. In particular, a radio map of the environment is first constructed through a specific offline training phase and is then used to estimate the target nodes' positions by best matching newly collected RSS values with those saved in the radio map. This technique implicitly takes into account the presence of reflections and multipath and is then particularly effective in indoor scenarios. Nevertheless, despite the appeal of fingerprinting for indoor localization applications, a few works have so far appeared in the literature.

Unlike existing works where fingerprinting is used to localize subjects in large (indoor) areas, in this work fingerprinting-based radio localization is exploited to perform arm posture recognition by estimating the positions of (i.e., localizing) target wireless sensor nodes properly placed on a user arm (i.e., body area localization). Different fingerprinting-based localization algorithms, either deterministic or probabilistic, are

*Corresponding author. E-mail: matteo.giuberti@xsens.com.
**Matteo Giuberti is with Xsens Technologies B.V. since April 2014. This work was performed while he was at the University of Parma.

considered to estimate targets' positions. Furthermore, the proposed radio-based posture recognition system is extended to integrate the use of inertial measurements. More precisely, measurements obtained from accelerometers (available on the target nodes) are used to estimate the pitch of every node in order to refine the accuracy of the position estimation provided by the fingerprinting-based radio localization.

We remark that the key contribution of this work mainly consists in experimentally investigating the feasibility of a hybrid approach which combines radio localization and inertial signals.[1] In particular, to the best of our knowledge, besides the novelty of the use of a radio localization technique (i.e., fingerprinting) for posture recognition applications, no work on the joint use of radio and inertial signals for these types of applications has appeared in the literature. Moreover, while in the literature inertial posture recognition approaches tend to use inertial devices equipped with accelerometers, gyroscopes, and/or magnetometers [7–9], our low-complexity approach relies only on acceleration measurements. The proposed system is extensively studied (optimizing key parameters) and its performance is also compared with that of a low-cost optical system (namely, Kinect). Finally, even if the proposed system is conceived to recognize simple postures of a static (not walking) user, the design and implementation of a fully portable ambulatory posture recognition system is an appealing research extension.

The rest of this paper is structured as follows. Section 2 is dedicated to related work. In Section 3, the proposed radio/accelerometric hybrid approach is presented and described, specifically focusing on the problem of arm posture recognition. In Section 4, the considered experimental set-up is presented and the performance metrics of interest are then introduced. In Section 5, the system performance, after parametric optimization, is analyzed. Finally, in Section 6, advantages, disadvantages, and possible future extensions of the proposed system are discussed.

2. Related work

In the literature, two main approaches have been proposed for posture recognition: optical and inertial.

Concerning *optical posture recognition*, the widely used technology is optoelectronic (e.g., Vicon system [10]). Optoelectronic systems require the user to wear reflective markers and to move in a space completely visible by a set of cameras. Because of their accuracy, these systems are typically used as ground truth reference for other posture recognition systems. On the other hand, their use is typically limited to clinical environments or specialized laboratories, due to their large cost and complexity.

Other optical posture recognition systems comprise the class of markerless systems. Kinect [11], which features an RGB camera and a depth sensor (composed by an infrared camera and projector), can be considered one of the most significant example of markerless systems. Its low cost (with respect to systems like Vicon), along with its still quite good performance, has made it a widely used solution. Nevertheless, it also has some spatial and temporal limitations, which could be critical in the context of some applications. More specifically, concerning spatial limitations, it is generally known that the z axis of Kinect, which is related to the direction perpendicular to its sensor camera, has poorer resolution with respect to those of its x and y axes, which instead define the frontal plane [12]. Moreover, Kinect's sampling rate, which is typically around 30 frames per second (fps), can be a limiting factor when monitoring fast movements.

Typically, optical systems suffer from problems related to different lighting conditions and markers' occlusion. Moreover, the user movements must be limited to the area captured by the cameras. The reader is referred to [13,14] for accurate surveys on optical posture recognition.

Concerning *inertial posture recognition*, inertial sensors are typically used to estimate the orientation of rigid body segments and, thus, to recognize the postures of a user. One of the most successful and complete commercial products is the Xsens Moven [15], which comprises 17 inertial sensors (equipped with tri-axial accelerometers, gyroscopes, and magnetometers) attached to the body of the user by a lycra suit. The major advantage of this technology, with respect to optical systems, is that the user is completely free to move everywhere because no camera is needed. Moreover, the visibility of the nodes placed on the user body is not an issue. However, the accuracy of these systems is typically lower than that of optical systems and the cost, particularly for systems which rely on a large number of nodes and types of sensors (such as Xsens Moven), is not significantly lower than that of optoelectronic systems. Finally, especially when used for a long time, a significant drift in the sensors' measurements can

[1] Preliminary results obtained with radio fingerprinting (without using inertial signals) can be found in [5,6].

be typically observed, leading to a performance degradation. It should also be observed that, unlike optical systems, inertial systems cannot *directly* provide information about the sensor nodes' positions. Instead, estimated sensor devices' orientations are used together with properly defined biomechanical models, comprising the lengths of all body segments, to reconstruct the full body posture of the user through forward kinematics techniques [16]. The user position is estimated using advanced contact detection techniques.

Finally, solutions based on the joint use of different technologies, designed in order to tackle and overcome the limitations which characterize component technologies considered independently, have also been widely investigated. For instance, inertial/GPS, inertial/optical, and inertial/acoustic joint measurements are considered in [17], in [18], and in [19], respectively. In particular: in the inertial/GPS approach, inertial nodes equipped also with GPS receivers are used to track the user, providing better performance (especially) in outdoor scenarios (e.g., sport sessions as skiing); in the inertial/optical approach, inertial sensor nodes and cameras are jointly used, leading to a decreased freedom of movement of the user but also to an improved robustness to occlusions; in the inertial/acoustic approach, nodes equipped with microphones and speakers are used in addition to inertial sensor nodes, providing better performance (thanks to the estimation of relative distances between nodes) but suffering from "acoustic occlusions."

Unlike the above approaches, our solution represents the first attempt (to the best of our knowledge) of investigating the effectiveness of the joint use of radio and inertial (specifically, accelerometric) measurements. To this end, an arm posture recognition system is developed in order to evaluate the feasibility of the use of such hybrid approach for posture recognition. Advantages and disadvantages are inherited from the considered technologies and our solution compares to existing ones as follows.

– *Optoelectronic systems* (e.g., Vicon) provide a better accuracy than that of our approach, at the price of (i) a much higher cost and (ii) problems related to different lighting conditions and markers' occlusion.
– *Markerless systems* (e.g., Kinect) suffer from problems similar to optoelectronic systems (i.e., related to different lighting conditions and markers' occlusion). They have a cost and a performance in the order of those of our approach. However, unlike our approach, a benefit of these systems is that no specific hardware has to be placed on the user body.
– *Inertial systems* have a performance which highly depends on the used inertial sensors. However, standard inertial systems (i.e., which use devices with accelerometers, gyroscopes, and magnetometers) provide better performance than that of ours and allow the user a higher freedom of movement, at the price of a higher cost.
– *Inertial/GPS* hybrid solutions have the same advantages/disadvantages of inertial systems, but requires necessarily to be outdoor. Of course, they are very suitable for outdoor posture recognition sessions (i.e., improved freedom of movement).
– *Inertial/optical* hybrid solutions, when considering devices with costs comparable to that of our system, are almost similar (in terms of advantages/disadvantages) to our approach.
– *Inertial/acoustic* hybrid solutions are also very similar to our approach. However, even if they suffer from problems related to "acoustic occlusions," they allow improved portability and freedom of movement.

Unlike our solution, all other systems do not need a training phase – more precisely, the radio-based component of our approach relies on a (short) training phase. However, preliminary calibration phases must be typically taken into account.

To summarize, the proposed radio/inertial hybrid approach is conceived as a low-cost and low-complexity approach which overcomes typical limitations of existing systems, such as those related to different lighting condition and occlusions. On the other hand, the proposed posture recognition system requires that the user does not walk (i.e., allows a limited freedom of movement to the user). To overcome these limitations, the design of an entirely portable posture recognition system will be discussed.

3. Arm posture recognition

In this work, "arm posture recognition" refers to the continuous estimation of the three-dimensional positions of sensor devices placed on the user arm and, therefore, to the continuous estimation of the arm orientation – this differs from "arm posture classification," where arm postures must be detected choosing from a discrete set of predefined postures. Indeed, even if our hybrid approach uses a localization technique (i.e., radio fingerprinting) to classify "known" (trained) arm postures (*posture classification*), it then properly interpolates these postures in order to estimate whole transitional movements of the arm (and not just "still postures") between two or more of the previous trained postures (*posture recognition*). Further considerations of the actual suitability and applicability of the proposed system for arm posture recognition (and classification) will be presented in Section 6.

3.1. Fingerprinting-based radio localization

As anticipated in Section 1, fingerprinting is a robust localization technique for indoor scenarios, which are typically characterized by reflections, multipath, and fading. We now provide some intuition on the fingerprinting technique – the interested user can find more details in [5,6]. Fingerprinting requires three kinds of nodes: target nodes, anchor nodes, and a base station. Target nodes have to be localized, whereas anchors nodes have fixed known positions and are used to generate a reference system. Finally, the base station is the processing center. Two phases are considered: a training phase, during which a radio map of fingerprints is generated, and an online phase, during which localization is performed.

During the training phase, the target node continuously broadcasts packets to be received by the anchor nodes. The latter, upon reception of the packets sent by the target, measure the RSSs and relay this information to the base station. The base station collects the RSS values and groups them into "fingerprint vectors." Finally, the base station generates a "fingerprint," i.e., a vector containing the arithmetic average of the received fingerprint vectors (i.e., the vector whose elements are the average RSSs measured by the anchors). Different physical positions of the target (i.e., "fingerprint positions") in the monitored area (properly chosen depending on the considered application) correspond and lead to different fingerprints. The entire set of fingerprints created in the training phase represents the "radio map" of the environment.[2] This radio map can be then used to run *deterministic* (i.e., based on simple comparisons between newly measured fingerprint vectors and the fingerprints of the radio map) localization algorithms. Furthermore, if one wants to use *probabilistic* versions of these algorithms (i.e., based on a more accurate statistical characterization of the RSS), during the training phase the entire Probability Mass Functions (PMFs) of the RSSs from all anchors need to be also computed and stored. Even if in the context of indoor localization a log-normal distribution of the RSS seems to be widely accepted [3], this is generally not our case, especially due to the fact that sensor nodes are placed on a human body. Therefore, the PMFs will be experimentally evaluated using the so-called histogram method, where the normalized histogram of the actual RSS measurements (during the training phase) for each fingerprint position are used [20].

After the training phase is completed, the online phase starts (following the same operations of the training phase, but now building a so-called "online vector," i.e., a single fingerprint vector containing a time snapshot of newly measured RSS values). In particular, in the online phase the radio map (deterministic approach) or the PMFs (probabilistic approach) created in the training phase are used to localize the target. Note that, once the training phase is over (i.e., the online phase starts), the target node can move freely and should not necessarily be placed in the previously trained fingerprint positions.

Given the measured online vectors, different algorithms can be used to estimate the positions of the target node. One of the simplest *deterministic* fingerprinting algorithm is the Nearest Neighbor (NN) algorithm, whose generalization is known as kNN[3] [21] – we remark that "neighbors" here refer to fingerprints and are thus associated with specific fingerprint positions. The kNN algorithm estimates the target positions by computing a specific distance metric between the online vector and every fingerprint contained in the radio map. By applying the Shepard method [22] in order to compute a weighed interpolation of the closest neighbors, the estimated target position $\hat{s} = (x, y, z)$ is given by

$$\widehat{s} = \sum_{i=1}^{k} \frac{w_i}{\sum_{j=1}^{k} w_j} \cdot \widehat{s}_i \qquad (1)$$

[2] The major strength of fingerprinting consists of the fact that the fingerprints implicitly take into account the impact of reflections and multipath on the RSSs, i.e., this technique is "tailored" to the specific indoor environment. This makes fingerprinting virtually insensitive to indoor propagation limitations – provided that the propagation environment remains quasi-static.

[3] Note that the term kNN, in the context of radio fingerprinting, is used to indicate the identification of the k closest fingerprints. Therefore, unlike in typical machine learning scenarios, the NN search is performed in a database of a very limited size (i.e., the number of rows is equal to the number of trained fingerprint positions) and, therefore, is not computationally intense.

where $\{\widehat{s}_i\}_{i=1}^k$ are the fingerprint positions (i.e., physical positions) of the k closest neighbors (i.e., the fingerprints with shortest distances from the online vector) and

$$w_i \triangleq \frac{1}{d_i^{p_s} + 0.0001} \tag{2}$$

where: d_i is the distance computed between the ith closest neighbor and the online vector (defined in the space of RSS vectors); and p_s is an integer larger than 0. The term 0.0001 at the denominator of Eq. (2) is used to prevent a division by zero if the online vector is equal to one of the fingerprints. In our system, we will consider only two definitions of distance: Euclidean and Manhattan [21]. Other distance definitions can be applied to the kNN algorithm, e.g., the Mahalanobis distance, which takes also into account the contribution of covariance matrix computed for every fingerprint [21] – little performance differences are, however, observed. Observe that, when $k = 1$ (i.e., with the NN algorithm), Eq. (1) reduces to the coordinates of the closest fingerprint position and, then, p_s has no influence on the system. Finally, note that, due to the interpolation between fingerprint positions in Eq. (1), the estimated position $\widehat{s}$ may (likely) differ from any of the considered fingerprint positions.

Unlike the deterministic approach, in the *probabilistic* approach (straightforwardly called p-kNN) the RSSs measured at the anchor nodes are characterized, using the samples received in the training phase, through their entire PMFs. In this case, the estimated target position can be expressed as follows:

$$\widehat{s} = \sum_{i=1}^{k} \frac{P(\widehat{s}_i|\boldsymbol{r})}{\sum_{j=1}^{k} P(\widehat{s}_j|\boldsymbol{r})} \widehat{s}_i, \tag{3}$$

where: $\boldsymbol{r}$ is the "online" vector; $P(\widehat{s}_i|\boldsymbol{r})$ is the a-posteriori probability of the ith (out of k) closest neighbor; and the k closest neighbors are chosen so that the corresponding a-posteriori probability is maximized. More specifically, using Bayes theorem, the a-posteriori probability that the target node is in the ith fingerprint position, given that the online vector $\boldsymbol{r}$ is received, can be expressed as

$$P(\widehat{s}_i|\boldsymbol{r}) = \frac{P(\boldsymbol{r}|\widehat{s}_i)P(\widehat{s}_i)}{P(\boldsymbol{r})} = \frac{P(\boldsymbol{r}|\widehat{s}_i)P(\widehat{s}_i)}{\sum_{\ell=1}^{L} P(\boldsymbol{r}|\widehat{s}_\ell)P(\widehat{s}_\ell)} \tag{4}$$

where: L is the number of trained fingerprint positions; $P(\boldsymbol{r}|\widehat{s}_i)$ is computed (as anticipated earlier in this subsection) using the histogram method (with a bin resolution of 1 dBm); and $P(\widehat{s}_i)$ is the a-priori probability of being in the ith fingerprint position. Since, with no movement restriction, all fingerprint positions are equally likely, it holds that $P(\widehat{s}_i) = 1/L$. Note again that, when $k = 1$ (and thus reducing to a so-called p-NN), Eq. (3) returns exactly the coordinates of the closest fingerprint position.

Concerning arm posture recognition, fingerprinting can be used to estimate the positions of (i.e., localize) multiple target nodes, properly placed on a user arm (e.g., one on the upper arm and one on the forearm), and straightforwardly derive the arm posture. To this end, the reference system origin must be properly chosen (e.g., the shoulder) and the user must try to keep this origin fixed during the evaluation. In this application scenario, the fingerprint positions correspond to targets' physical positions related to predefined arm postures that must be held by the user during the training phase. During the following online phase, the user can instead move his/her arm freely (always recalling not to move the reference system origin, i.e., his/her shoulder).

3.2. *Accelerometer-based pitch estimation*

The radio fingerprinting-based posture recognition system described in Section 3.1 may introduce errors, especially when a target node is in a position that differs from the trained fingerprint positions (i.e., when the arm is in an untrained posture). A possible way to improve the system performance is to estimate the arm orientation by making use of other inertial sensors (e.g., accelerometers, gyroscopes, and/or magnetometers), which the target nodes can be equipped with. In particular, considering proper combinations of these sensors (e.g., an accelerometer and a gyroscope), the orientation of a device (and, thus, of the arm) can be estimated [7].

The orientation of a device can be described by three parameters: yaw (or heading), pitch (or elevation), and roll (or bank) [23]. Specifically, it is known that a rigid body can be arbitrarily rotated by first rotating it around its z axis by an angle ψ (the yaw), then around its y axis by an angle θ (the pitch), and finally around its x axis by an angle ϕ (the roll) [23]. Observing that the acceleration measured by a still device is only due to the gravity acceleration, it can be shown that, using just an accelerometer (in order to minimize the cost of

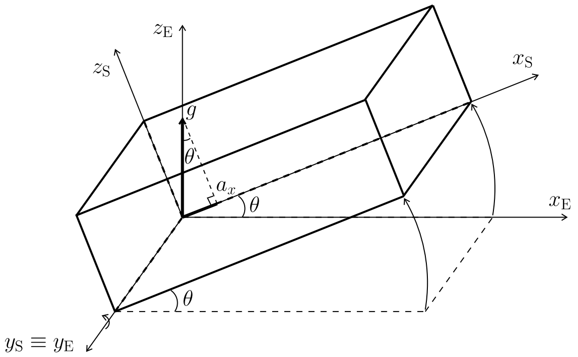


Fig. 1. Measured acceleration when the device is rotated around its y axis (i.e., its pitch θ is different than 0). Both Earth (E subscript of axes) and device (S subscript of axes) coordinate systems are shown.

the system), the pitch of a *still* device (i.e., following the previous notation, the angle between its x axis and the horizontal plane perpendicular to the gravity direction) can be determined by observing how the gravity vector is rotated with respect to the x axis of the device. More precisely, by exploiting acceleration measurements and following simple trigonometric equations, the device pitch θ can be computed as follows [24]:

$$\theta = \arcsin \frac{a_x}{g} = \arcsin a_x \tag{5}$$

where a_x is the acceleration (in g units) measured along the x axis and g is the gravity acceleration (obviously, equal to 1 g). For ease of clarity, in Fig. 1 a graphically intuitive representation of the geometrical meaning of Eq. (5) is provided. Note that Eq. (5) holds since the device is still and, therefore, the measured acceleration vector $\boldsymbol{a}$ has the same direction and norm of the gravity vector $\boldsymbol{g}$.

As the above approach is valid only if the device is still, static (constant force of gravity) and dynamic (movements or vibrations of the accelerometer itself) accelerations need to be discriminated in the case the device is moving. This problem cannot be easily solved, but it can be at least mitigated by taking into account only acceleration measurements with amplitudes in $[g - \xi, g + \xi]$, where g is the gravity acceleration (i.e., 9.81 m/s^2) and ξ needs to be properly chosen with respect to the application context. This can be carried out by considering only data portions in which the user is not moving (and, thus, $\xi \cong 0$ m/s^2) or by simply applying a low-pass filter to the output acceleration signals.

When the device is moving, since it is no longer true that $|\boldsymbol{a}| = g = 1$ g and since the arcsin function accepts only values whose norm is equal to or lower than 1, Eq. (5) should be properly rearranged. To this end, the estimated pitch of the device can be estimated as follows:

$$\theta = \arcsin \overline{a}_x = \arcsin \frac{a_x}{|\boldsymbol{a}|} \tag{6}$$

where $\overline{a}_x$ is the normalized acceleration (in g units) measured along the x axis.

Similarly, when the sensor device is attached to a rigid body segment of the user (e.g., the arm), the pitch (i.e., the inclination) of the considered body segment can be computed as

$$\theta = \arcsin \overline{a}_{\mathrm{bs}} \tag{7}$$

where $\overline{a}_{\mathrm{bs}}$ is now the normalized acceleration (in g units) measured along the device axis aligned with that of the considered body segment. Note that, according to Eqs (5), (6), and (7), the pitch always belongs to $[-\pi/2, \pi/2]$.

Since the estimated inclination of a body segment is not sufficient alone to provide information about its position, a proper human biomechanical model, which assumes to know the involved body segments' lengths and how they are linked together, should be considered.

3.3. Recursive estimation process

Referring to the posture recognition of a user arm, it will be now described how estimates obtained from (i) radio fingerprinting and (ii) pitch estimation should be properly combined. In particular:

- The pitch θ of each device (and, thus, of the corresponding body segment) is computed using Eq. (7). This will lead to the estimation of the z coordinate of the target node.
- Estimates of the (x, y, z) coordinates of the devices are available from radio localization (through fingerprinting).

Furthermore, we preliminary introduce a few geometric assumptions.

- A (properly defined) fixed origin **O** is chosen.
- One device is attached to each adjacent rigid body segment between the previously chosen origin **O** and the farthest distal point **Z** of the body part whose posture needs to be estimated. More precisely, for arm posture recognition, **O** could be the shoulder and **Z** could be the wrist. Therefore, two

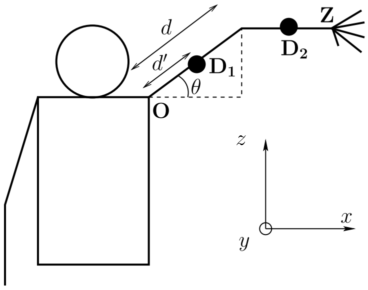


Fig. 2. Pictorial description of the considered scenario. The points of interest used in the algorithm are highlighted. In particular: **O** and **Z** are the 3D coordinates of the origin point and the farthest distal point that needs to be estimated, respectively; $\mathbf{D}_1$ and $\mathbf{D}_2$ are the 3D coordinates of target 1 and target 2, respectively; d is the actual length of the arm segment considered at the first step (in this case, the upper arm); d' is the distance between the considered sensor device (with coordinates $\mathbf{D}_1$) and the corresponding body segment's joint proximal to the fixed origin (in this case, the origin point with coordinates **O**, i.e., the shoulder); and θ is the pitch of the considered sensor device and, thus, of the corresponding body segment (in this case, the upper arm).

devices should be used, attached, respectively, to the upper arm and to the forearm.

- The length d of each body segment on which the devices are mounted is known (e.g., the lengths of the upper arm and of the forearm).
- The distance d' between each device and the corresponding body segment's joint proximal to the fixed origin is known (e.g., the distance between the shoulder and the device attached to the upper arm).

In Fig. 2, a pictorial description of the scenario, with highlighted points of interest, is shown.

In the following, we introduce a recursive estimation process, according to which adjacent body segments are considered at consecutive steps. Specifically, all adjacent body segments comprised between **O** and **Z** are considered, starting from the body segment starting in **O** and ending with the body segment ending in **Z**. Furthermore, for each considered body segment, the positions of its two extremes (i.e., the body segment joints) are chosen as reference points: the position of the proximal joint (i.e., the joint closest to the body) is denoted as point $\mathbf{A} = (x_\mathrm{A}, y_\mathrm{A}, z_\mathrm{A})$ and the position of the distal joint (i.e., the joint farthest from the body) as point $\mathbf{B} = (x_\mathrm{B}, y_\mathrm{B}, z_\mathrm{B})$. In the specific case of the arm shown in Fig. 2, at the first step **A** corresponds to **O**, whereas at the last step **B** corresponds to **Z**.

In order to estimate the posture of the user arm, the postures of the upper arm and of the forearm have to be estimated (or, in other words, the positions of the elbow and the wrist, with respect to the shoulder, need to be identified). In this case, the proposed recursive estimation strategy involves two steps. To this end, two target nodes need to be used: the first on the upper arm and the second on the forearm – more details on the experimental set-up will be given in Section 4.1. For ease of clarity, let us denote their physical positions as $\{\mathbf{D}_i\}_{i=1}^2$, where $\mathbf{D}_i = (x_{\mathrm{D}_i}, y_{\mathrm{D}_i}, z_{\mathrm{D}_i})$.[4] The origin $\mathbf{O} = (0, 0, 0)$ will be the shoulder, whereas the farthest distal point **Z** will be the wrist.

As shown in Fig. 2, at the first step, since we are focusing on the upper arm, **A** corresponds to the shoulder (i.e., the origin **O**), whereas **B** corresponds to the elbow (whose physical position is currently unknown). We then consider the first device, denoted as point $\mathbf{D}_1$ and (in our scenario) positioned in the middle of the upper arm, i.e., between **A** and **B**. Its initial position estimate, denoted as $\widehat{\mathbf{D}}_1$, is recovered through radio localization, whereas its pitch θ (which corresponds to the inclination of the upper arm) is estimated using Eq. (7). Finally, d is the distance between **A** and $\mathbf{D}_1$, whereas d' is the length of the upper arm (i.e., the distance between **A** and **B**).

Taking into account a polar coordinates framework, the arm posture is then estimated through the computation of the pitch and the heading of all the body segments of the arm. In particular, the main idea is to use the initial position estimate $\widehat{\mathbf{D}}_1$ (recovered from radio localization) to infer information about the heading ψ of the considered body segment. The heading information is then combined with that of the pitch θ of the same body segment (estimated using the accelerometer) in order to refine and correct $\widehat{\mathbf{D}}_1$ (taking also into account a proper biomechanical model). Therefore, the hybrid approach consists in using the (i) radio fingerprinting to estimate the heading of each arm segment and the (ii) inertial signals to estimate the inclination.

More precisely, the heading ψ of the considered body segment can be computed as

$$\psi = \arctan 2(y_{\widehat{\mathrm{D}}_1} - y_\mathrm{A}, x_{\widehat{\mathrm{D}}_1} - x_\mathrm{A}) \tag{8}$$

where the function arctan 2 behaves as the standard arctan function but, in addition, provides information

[4]Note that we assume that the z axis of the reference system is in the vertical direction (i.e., that of the gravity vector). The x and y axes may be instead conveniently chosen as perpendicular axes in the horizontal plane. In particular, we will chose the y axis as the forward direction and the x axis as the direction normal to the user sagittal plane.

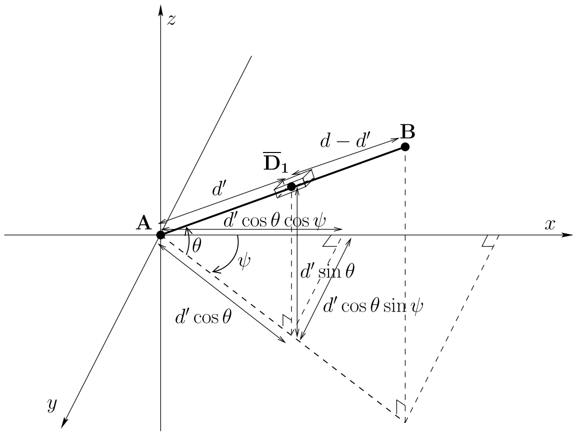


Fig. 3. Graphical intuition of Eq. (9) in the specific case of the first considered body segment (i.e., the upper arm). The heading (i.e., ψ) and the pitch (i.e., θ) of the segment are highlighted. For ease of clarity, the device $\mathbf{D}_1$ is pictured along segment **AB** and the projection of the segment on the horizontal plane is also shown. Note that, for the case of the upper arm, **A** and **B** correspond, respectively, the shoulder and the elbow of the user.

about the quadrant of the computed angle, so that the heading ψ belongs to $[0, 2\pi]$.

Finally, considering the estimated pitch (θ) and heading (ψ), the adjusted estimated coordinates of the first device, denoted as $\overline{\mathbf{D}}_1$, can be expressed as

$$\overline{\mathbf{D}}_1^T = \begin{pmatrix} x_{\overline{\mathrm{D}}_1} \\ y_{\overline{\mathrm{D}}_1} \\ z_{\overline{\mathrm{D}}_1} \end{pmatrix} = \begin{pmatrix} x_\mathrm{A} + d' \cdot \cos\theta \cdot \cos\psi \\ y_\mathrm{A} + d' \cdot \cos\theta \cdot \sin\psi \\ z_\mathrm{A} + d' \sin\theta \end{pmatrix} \quad (9)$$

where $(\cdot)^T$ is the transpose operator. A graphical intuition of Eq. (9) is given in Fig. 3. Furthermore, the estimated coordinates of **B** (namely, the elbow) can be similarly expressed as follows:

$$\mathbf{B}^T = \begin{pmatrix} x_\mathrm{A} + d \cdot \cos\theta \cdot \cos\psi \\ y_\mathrm{A} + d \cdot \cos\theta \cdot \sin\psi \\ z_\mathrm{A} + d \sin\theta \end{pmatrix}. \quad (10)$$

The above estimation process can be repeated to estimate the position of the second device $\mathbf{D}_2$, positioned in the middle of the forearm (as shown in Fig. 2), and of the wrist. To this end, **A** and **B** should be updated (i.e., **A** would correspond to the elbow, whereas **B** would correspond to the wrist) and Eqs (9) and (10) can be used again considering proper values of θ, ψ, d', and d (namely, the ones related to the arm segment between the updated points **A** and **B**, i.e., the forearm). Finally, since after this second step the updated point **B** (i.e., the wrist) corresponds to **Z**, the recursive process ends and the arm posture is estimated.[5] Note that as the position of **B** is estimated and, therefore, may have errors, it is likely that the estimated coordinates of $\mathbf{D}_2$ will be less accurate than those of $\mathbf{D}_1$.

Finally, we remark that the integration of the accelerometers into our system does not have any impact on the fingerprinting training phase. Indeed, all the acceleration measurements are taken into account only during the online phase.

4. Experimental set-up

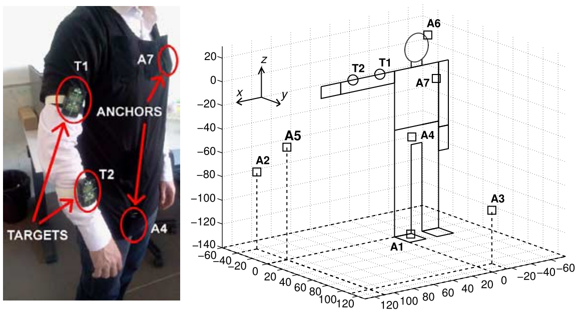
4.1. Experimental testbed

SunSPOT devices have been used for the experimental testbed. SunSPOTs are wireless devices equipped with a triaxial accelerometer and an IEEE 802.15.4 compliant radio interface with an on-board antenna and up to 100 m transmission range [25].

In the current paper, we consider a testbed which extends the one proposed in [5]. In particular, a Body Area Network (BAN) with target nodes on the user arm is still considered, but now the anchor nodes are in part placed (and fixed) in the surroundings of the user and in part attached on his/her body – in [5], all anchors are fixed and outside the body. This is obtained through the use of (i) a home-made t-shirt with folders where anchors can be placed and (ii) a hat with an anchor attached to it. Although the presence of some anchor nodes at fixed (outside the body) positions still forces the user to remain in its initial position in the room (in order to make the fingerprinting technique work consistently), the proposed testbed is a first step toward a fully portable arm posture recognition system – this extension is the subject of our current research activity. Furthermore, a noisier (and, thus, more realistic) environment is here taken into account (e.g., with more than one people moving in the room while testing the system and in the presence of multiple active WiFi networks), whereas in [5,6] the system performance is analyzed in more controlled (interference-free) scenarios. Therefore, the presented results are already reflecting the performance of a system in a quite adverse (but realistic) home scenario.

In the present experimental testbed, 2 SunSPOTs, acting as targets (i.e., target 1 and target 2), are placed on the right arm of the user, as shown in Fig. 4(a)

[5]Observe that this process can be recursively repeated, should more consecutive body segments be considered (e.g., by adding a device for the hand in our testbed).



(a) (b)

Fig. 4. Considered experimental posture recognition set-up: (a) plot of on-body set-up; (b) scheme of the overall set-up. The positions of the anchor nodes (i.e., A1 to A7) and of the target nodes (i.e., T1 and T2) are highlighted. Note that A2, A3, and A5 are fixed in the surroundings of the user, whereas A1, A4, A6, and A7 are attached on his/her body.

Table 1

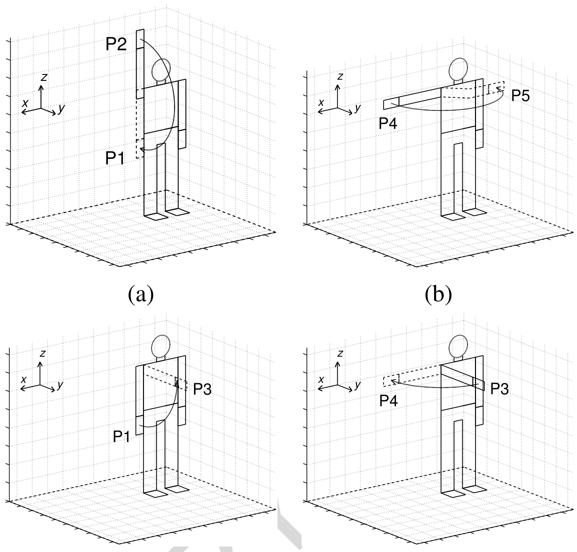
Fingerprint positions coordinates (considered during the training phase) with respect to the user shoulder. Every fingerprint position corresponds to a specific arm posture

Fingerprint Positions	Target 1			Target 2		
	x [cm]	y [cm]	z [cm]	x [cm]	y [cm]	z [cm]
P1	0	0	−15	0	0	−42.5
P2	0	0	15	0	0	42.5
P3	0	15	0	0	42.5	0
P4	15	0	0	42.5	0	0
P5	−11	11	0	−34	21	0

(where it is also possible to see the home-made t-shirt): the first node (i.e., target 1) is on the upper arm and the second (i.e., target 2) is on the forearm. $N = 7$ anchor nodes are considered: 4 of them are placed on the user body, whereas the remaining 3 are placed in its proximity.[6] For ease of clarity, an illustrative representation of the overall experimental set-up is shown in Fig. 4(b), where a map of the positions of the nodes is shown. Fixing the origin of our new reference system on the shoulder (and thus forcing the user to move the arm and keep the shoulder still), 5 fingerprint positions per target node (i.e., P1, P2, P3, P4, and P5), whose coordinates are defined in Table 1 and shown in Fig. 5, are considered, corresponding to 5 simple arm postures. The choice of the 5 arm postures (and, thus, of the 5 fingerprint positions) is expedient to cover as uniformly as possible the surroundings of the user arm.

The packet transmission rate of target nodes is around 30 pck/s. The acceleration, locally (at each target node) sampled at 100 Hz, is also consistently down-sampled to the same rate (i.e., 30 Hz). Therefore, every second, the base station receives, for each target, around 30 new fingerprint vectors (containing new RSS measurements at anchor nodes) and new acceleration measurements. The synchronization between the sensor nodes in the system is performed by aligning the internal clocks of every sensor node to that of the base station, at application deployment stage, and by inserting a timestamp field inside each transmitted packet.[7]

In the training phase, the user has been asked to keep each of the 5 arm postures for about 30 s. As shown in [6], the time interval, during which each arm posture should be kept fixed to train each corresponding fingerprint position (i.e., 30 s), has been chosen in order to collect a number of fingerprint vectors (i.e.,



(a) (b) (c) (d)

Fig. 5. Transitional movements, performed by the user during the online phase, between the 5 trained arm postures (used to train the corresponding 5 fingerprint positions, i.e., P1, P2, P3, P4, and P5) considered in the training phase.

[6]We remark that, in the presented experimental analysis, the user is forced to keep his/her body still. Therefore, the anchor nodes attached on his/her body can be actually considered as fixed (still) nodes. The choice of attaching them directly on the user body will be further discussed and motivated in Section 6.

[7]Note that, even if for the purpose of our experimental analysis this is sufficient to guarantee a fair synchronization for a few hours, a future system should try to define specific communication protocols/mechanisms able to achieve, e.g., through a sparse exchange of synchronization packets, a long-lasting synchronization of system devices.

about 1000 for each target node) sufficient for the convergence of the evaluated fingerprints. During these 30 s, the considered fingerprint position is trained as explained in Section 3.1 and the related fingerprint is generated.[8] Once the training phase has terminated, the online phase is split into two parts in order to properly test the system performance. In particular: the user is first asked to replicate the 5 trained arm postures (in order to test the system in *static* conditions); furthermore, a few transitional movements, which start and end at two of the trained arm postures, are also executed (in order to test the system in *dynamic* conditions). Note that, for the purpose of localization, fingerprinting is performed independently for each target node. With reference to Table 1, this means that: P1 corresponds to a specific posture of the arm but, obviously, to two fingerprint positions (i.e., one for each target node) and, therefore, to two fingerprints; P2–P5 can be interpreted likewise. The transitional movements that the user is asked to perform are shown in Fig. 5 and can be summarized as follows:

- lower the arm from position P2 to position P1, passing through position P3 (Fig. 5(a));
- move horizontally the arm from position P4 to position P5, passing through position P3 (Fig. 5(b));
- raise the arm from position P1 to position P3 (Fig. 5(c));
- move horizontally the arm from position P3 to position P4 (Fig. 5(d)).

The previous sequence of transitional movements is repeated by the user twice, resulting in a total of 8 movements. The user is asked to perform each transitional movement at constant speed (typically leading to movements with duration of $2\div4$ s). Since the segmentation of the time portions containing each transitional movement is only expedient to the performance analysis (and not to the correct system behavior), it has been performed manually. During the online phase, each new fingerprint vector (i.e., the online vector) is used to evaluate the current arm posture.

Finally, for the purpose of a *trend-wise* comparison, we run at the same time an arm posture recognition session performed through the Kinect system [11]. In particular, we put the Kinect device in front of the user (at a distance of about 3 m) and run the Skeletal Viewer demo application (included in the Kinect libraries). Thank to this application, we were able to save the coordinates of the joints of the user arm for each time sample and, therefore, the coordinates of the target nodes, which have been compared with the ones estimated with our system. Given that Kinect samples video frames at around 30 fps, a direct comparison with our system is straightforward. Five users have been considered for the experimental analysis.

[8] Note that, even if some slight variations of the kept arm postures are unavoidable, the measured fingerprint vectors should be relatively close to each other (in the RSS space). Therefore, the error introduced in the computation of the fingerprints is typically negligible.

4.2. *Performance metrics*

Typically, the performance of an arm posture recognition system is evaluated by comparing the estimated arm position with that estimated with optoelectronic systems (e.g., Vicon) and considered as “ground truth.” For ease of clarity, we remark that our goal is not to derive the *exact* performance of the proposed system, but, rather, to have a rough idea of its achievable performance, in a comparative way with respect to Kinect. Since we do not have access to any optoelectronic system and since we are interested on *trend-wise meaningful* results, the performance of our system has been evaluated with respect to predefined arm postures and trajectories of the considered body segments, where true target positions have been manually derived – we believe that this is reasonable, as the considered postures and transitions (described in Section 4.1) are very regular. Due to the geometry of the human body (and, specifically, the arm model) and the simplicity of the chosen and tested transitional movements (described in Section 4.1), the considered trajectories are simple horizontal/vertical arcs in the three-dimensional space (as shown in Fig. 5). Each considered transitional movement starts and ends in correspondence to two fingerprint positions. Since the physical coordinates of the fingerprint positions are known, it is easy to also determine, with high accuracy, the coordinates of points lying on the arcs (namely, the continuous positions of the targets during the transitional movements). Furthermore, the users are asked to try to perform these movements at constant speed and, therefore, physical positions of intermediate points along the trajectories can be computed by sampling at constant rate the arcs’ points. Due to the simplicity of the considered movements, these trajectories can be accurately reproduced by a careful user. Even if, in this way, some errors are certainly introduced in the measurements, due to the simplicity of the chosen postures and movements these errors tend to be limited to a few centimeters. As introduced at the beginning of this subsection, we refer to our results as "trend-wise meaningful" to advise the reader that the use of a more precise ground truth (e.g., based on the use of an optoelectronic system) might lead to slight differences in the results. However, this limited discrepancy does not hinder the validity of the proposed framework and of the considered feasibility prototype. The performance of Kinect is evaluated by following the same experimental approach considered for the proposed system.

The performance of a localization algorithm can be characterized in terms of *accuracy* and *precision* [26]. In particular: *accuracy* is defined as the distance between the estimated position and the true target position; *precision* is the percentage of successful position estimates within a given accuracy. For instance, a precision of 60% with an accuracy of 10 cm means that 60% of the estimation errors made by the system are lower than or equal to 10 cm. The performance of a localization algorithm can be then fully described by evaluating the precision as a function of the accuracy, i.e., through a curve. Note that the best achievable performance corresponds to the (0, 1) point in the precision/accuracy graph, whereas the performance decreases as much as the distance between the precision/accuracy curve and this point increases.

Though the accuracy/precision curve provides significant insights on the system performance, in order to provide a more concise (yet insightful) system performance metric, the position error (PE), averaged along the duration of the considered arm movements, is also evaluated. Practically, given a movement of duration T and starting at t_0, PE is computed as follows:

$$\begin{aligned}\mathrm{PE} &\triangleq \frac{\sum_{t=t_0}^{t_0+T-1} d_{\mathrm{E}}(\hat{\boldsymbol{s}}_t, \boldsymbol{s}_t)}{T} \\ &= \frac{\sum_{t=t_0}^{t_0+T-1} \sqrt{(\hat{x}_t - x_t)^2 + (\hat{y}_t - y_t)^2 + (\hat{z}_t - z_t)^2}}{T}\end{aligned} \tag{11}$$

where: $\hat{\boldsymbol{s}}_t = (\hat{x}_t, \hat{y}_t, \hat{z}_t)$ and $\boldsymbol{s}_t = (x_t, y_t, z_t)$ are the estimated and the "true" target positions, respectively, at the tth epoch; and d_{E} stands for Euclidean distance.

5. Results

The performance of the proposed posture recognition system has been evaluated in both static and dynamic conditions. Concerning the system performance in *static* conditions (not reported here for lack of space), our results show that the deterministic approach performs slightly better than the probabilistic one: the average PEs are 3 cm and 4.8 cm for the deterministic and the probabilistic approaches, respectively. This confirms previous results in [5,6].

We now focus on the performance of our system in *dynamic* conditions (i.e., when the user arm moves over time). To this end, each user is asked to perform the sequence of the four transitional movements defined in Section 4.1 and shown in Fig. 5. Moreover, for repeatability purposes, the same sequence is repeated twice (leading to a total of 8 transitional movements).

In order to evaluate the system performance, the system parameters used in Eqs (1) and (3) (i.e., the number of anchors N and the related optimal subset, the values of k and p_{s}, and the distance metrics) are optimized in order to minimize the PE between the real targets' positions and their estimates. In particular, the optimal system parameters are determined both (i) considering independently each distinct transitional movement (i.e., determining 8 set of optimal parameters, one per movement) and (ii) considering jointly all transitional movements (i.e., determining a unique set of optimal parameters to be used for every possible movement).[9] Note that the presented results are averaged over the five users. Therefore, the optimal parameters are chosen as the ones which jointly maximize the average performance of all the users (i.e., they are not optimized independently for each user). Both deterministic and probabilistic approaches are considered. Finally, the performance of Kinect is also evaluated.

The system performance obtained by *optimizing independently each distinct transitional movement* is shown in Table 2.[10] It is easy to see that the posture recognition system works at its best when the user

[9]Note that, for a practical implementation of the system, a unique set of optimal parameters must be considered. Nevertheless, the optimization of the system parameters performed considering independently each transitional movement allows to evaluate an insightful lower bound of the performance of our system.

[10]Observe that, given that the user performs twice the same sequence of 4 movements, the movements from 5 to 8 (as referred to in the "Transitional Movements" column) are repetitions of the same movements that are identified by the labels from 1 to 4. Therefore, it is expected that the performance (and optimized parameters) in correspondence to movement 1 will be quite similar to that in correspondence to movement 5 – the same comment applies to the pairs of movements 2–6, 3–7, and 4–8.

Table 2
Posture recognition performance (optimized independently for each distinct transitional movement)

Target Nodes	Transitional Movements	Deterministic Approach						Probabilistic Approach				Kinect
		PE [cm]	N	anchors subset	distance	k	p_s	PE [cm]	N	anchors subset	k	PE [cm]
Target 1	1	**1.2**	3	{1, 5, 6}	Manhattan	5	1	**2.6**	1	{5}	3	**8.4**
	2	**4**	3	{5, 6, 7}	Euclidean	2	2	**6.7**	2	{6, 7}	3	**7.1**
	3	**1.3**	4	{2, 4, 5, 6}	Manhattan	1	–	**1.3**	1	{1}	3	**7.5**
	4	**4.5**	4	{3, 4, 5, 7}	Manhattan	4	6	**7.4**	1	{4}	4	**11.4**
	5	**1.3**	3	{1, 5, 6}	Manhattan	5	1	**3.4**	1	{1}	3	**8.1**
	6	**4.2**	3	{5, 6, 7}	Euclidean	3	2	**6.3**	2	{6, 7}	3	**6.9**
	7	**0.4**	2	{2, 4}	Manhattan	1	–	**0.4**	1	{ 1}	3	**7.8**
	8	**4.4**	4	{1, 4, 5, 7}	Euclidean	3	2	**6.3**	2	{6, 7}	2	**7.7**
Target 2	1	**3.3**	2	{2, 3}	Manhattan	1	–	**10.7**	2	{2, 7}	1	**18.1**
	2	**15**	4	{1, 2, 3, 6}	Manhattan	1	–	**21.9**	2	{3, 4}	1	**12**
	3	**3.6**	3	{3, 4, 7}	Euclidean	5	4	**4.5**	2	{4, 7}	1	**15.6**
	4	**12.5**	6	{1, 2, 3, 4, 5, 6}	Euclidean	2	6	**19.9**	4	{1, 2, 3, 4}	2	**25.9**
	5	**3.6**	2	{2, 3}	Euclidean	1	–	**9.4**	3	{ 2, 3, 7}	1	**17.2**
	6	**13.9**	3	{2, 3, 6}	Euclidean	1	–	**22.8**	4	{2, 5, 6, 7}	1	**12.5**
	7	**1.2**	3	{3, 4, 7}	Euclidean	1	–	**2.4**	3	{3, 4, 7}	3	**17.1**
	8	**12.7**	5	{3, 4, 5, 6, 7}	Manhattan	2	4	**15.1**	4	{2, 4, 6, 7}	2	**17.6**
Target 1 + Target 2	1÷8	**average PE = 5.4 cm**						**average PE = 8.8 cm**				**average PE = 12.6 cm**

performs movements 1 (or 5) and 3 (or 7). This provides insightful information about the system behavior which can be easily understood by observing the nature of these movements. Indeed, these movements are all "vertical" movements (i.e., the majority of the movements of the arm is concentrated on a vertical plane), whereas the others are all "horizontal" movements (i.e., the majority of the movements of the arm is concentrated on a horizontal plane). Taking into account our system, it is easy to notice that the "vertical" movements benefit more from pitch estimation (performed through the accelerometers) than from heading estimation (which derives entirely from the radio localization). On the other hand, the opposite comment stands for the "horizontal" movements, where the heading of the arm varies typically more than its pitch. This is then a strong symptom of the importance of the role of the accelerometers in our system.[11] As expected, in Table 2, it can also be observed that the nature of the movements (i.e., "vertical" or "horizontal") has no influence on the Kinect performance (which is reported in the last column of the table). Finally, in order to provide a general and concise indication of the system performance, the last row of the table reports the average PE (over all movements). It can be noticed that, in terms of average PE, our system slightly outperforms Kinect. Nevertheless, we remark that the comparison with Kinect is only trendwise meaningful and, thus, a limited effort has been dedicated to arrange the best acquisition conditions (for example, considering more favorable light conditions or optimized user's distance for Kinect). However, no significant performance improvement is expected with Kinect. In fact, previous investigations on the purpose of Kinect, in similar application scenarios, have also shown the presence of significant error peaks (even higher than 10÷15 cm) associated with similar arm movements [12,27]. As already anticipated, since it is impractical to tune the system differently for each movement, we remark that the average results in Table 2 are not directly applicable. Nevertheless, they can be used as "best-case" benchmark values and, therefore, to lower bound the performance of the posture recognition system.

Considering now a *global* (more practical) *optimization* computed upon the whole sequence of movements (i.e., by selecting a single set of parameters for all movements), the obtained system performance is shown in Table 3. As expected, the average PE is higher than in the previous case. In particular, it can be

[11] In order to improve the accuracy of estimation of "horizontal" movements the use of a gyroscope and/or a magnetometer is expedient.

Table 3
Posture recognition performance (optimized on the whole sequence of movements)

Target Nodes	Deterministic Approach						Probabilistic Approach				Kinect
	PE [cm]	N	anchors subset	distance	k	p_s	PE [cm]	N	anchors subset	k	PE [cm]
Target 1	**4.2**	4	{3, 5, 6, 7}	Euclidean	4	2	**6.1**	2	{6, 7}	5	**8.1**
Target 2	**14**	3	{2, 3, 6}	Euclidean	1	–	**17.9**	4	{2, 3, 4, 7}	1	**17.1**
Target 1 + Target 2	**average PE = 9.1 cm**						**average PE = 12 cm**				**average PE = 12.6 cm**

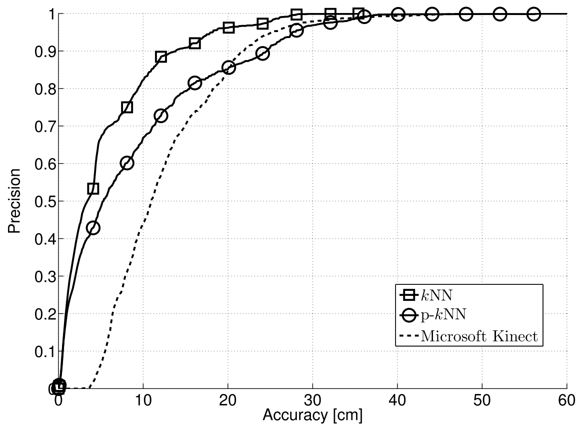
seen that the proposed arm posture recognition system guarantees an average PE around 10 cm. Specifically, the deterministic approach (with average PE equal to 9.1 cm) performs better than the probabilistic approach (whose corresponding average PE is 12 cm). As before, our system slightly outperforms the Kinect system (whose performance is again indicated in the last column of the table). The fact that the deterministic approach outperforms the probabilistic one (even if the opposite is intuitively expected) is likely due to the fact that the probabilistic version of the proposed system may need a longer training time or, at least, a larger number of training samples. In fact, we have observed that some RSS PMFs (determined during the training phase) have almost zero variance, which implies a lower flexibility (in the online phase) to handle small variations and noise in the measured RSS values. To this end, other solutions relying on more sophisticated methods, such as non-parametric estimation (e.g., Parzen-window density estimation [21]) or generative models (e.g., Gaussian Mixture Models [20]), which are more effective in the presence of data with small variance and correlation, may be also taken into account.

For ease of clarity, note that the optimal parameters indicated in Table 2 and Table 3 often differ, especially when considering different target nodes. From a practical applicability point of view (i.e., in terms of a final complete prototype), this does not mean that different configurations of anchor nodes have to be used for each target node. In fact, one should consider a single configuration given by the union of all anchor nodes present in all configurations, i.e., one should use the smallest set of anchor nodes which contains, as subsets, the optimal configurations for each movement of each target (e.g., {2, 3, 5, 6, 7} for the deterministic approach in Table 3). It is then the base station that, upon collection of all RSS data, in order to estimate the position of a target node, needs to consider only the RSS measurements obtained from the anchor nodes associated with the corresponding optimal subset. In fact, the use of an original (redundant) configuration of anchor nodes was just expedient to the investigation of the most suitable positions (and numbers) of anchor nodes in the surroundings of the user.

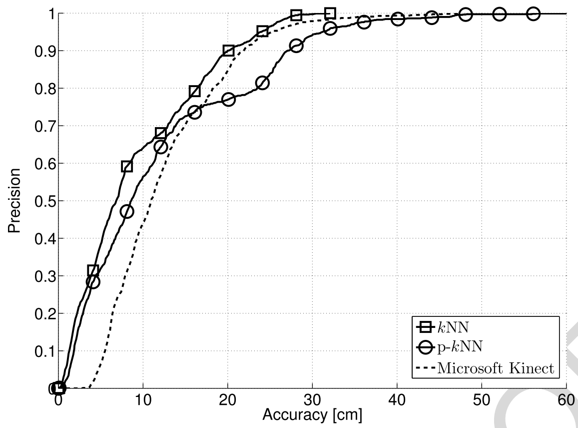
In order to better investigate the behavior of the proposed system, in Fig. 6 we finally evaluate the accuracy/precision performance of the proposed system. Both deterministic and probabilistic approaches are considered. The Kinect performance is also shown for comparison purposes. In Fig. 6(a) (which is related to results shown in Table 2), the parameters are optimized independently for each distinct transitional movement and, therefore, the obtained performance corresponds to a best-case scenario, whereas, in Fig. 6(b) (which is related to results shown in Table 3), the parameters are optimized considering jointly all possible transitional movements. As already observed in Tables 2 and 3, it can be observed that, in both cases (a) and (b), the deterministic approach slightly outperforms the probabilistic one and the Kinect system. Note that, in case (b), the Kinect system outperforms the probabilistic version of our system for accuracy values higher than 15 cm (in other words: if a position error larger than 15 cm is tolerable the Kinect system is more suitable than the probabilistic approach of our system).

6. Discussion and conclusion

As already anticipated throughout the paper, the main goal of this work is to investigate the feasibility of the combination of two known technologies, i.e., *radio (fingerprinting-based) localization* and *inertial (accelerometric) measurements*. Specifically, the efficacy of the considered hybrid radio/inertial approach in recognizing arm postures and movements has been evaluated. Our experimental analysis has been carried out using a SunSPOT testbed. Nevertheless, the proposed approach is more general and can be applied to any "node" equipped with an accelerometer and a radio interface.



(a)



(b)

Fig. 6. Precision, as a function of the accuracy, for the arm posture recognition experimental testbed, considering both deterministic and probabilistic approaches. The optimal configuration of the parameters is considered for every curve. In particular, in (a) every distinct transitional movement is optimized independently, whereas in (b) the optimization is carried out for the whole sequence of transitional movements. A comparison with the Kinect performance is also offered.

As previously discussed in Section 2, most of current state-of-the-art solutions for arm posture recognition rely on optical and/or inertial approaches. In particular, optical approaches, by using costly systems composed by a set of cameras and reflective markers placed on the user, allow to obtain an exceptional accuracy, but suffer from problems related to different lighting conditions and markers occlusions. On the other hand, these limitations are effectively tackled by inertial systems, which do not depend on cameras and, thus, allow an improved freedom of movement and portability, at the price of a lower accuracy. Inertial posture recognition is typically based on the estimation of the orientations of wireless devices placed on the body of a user, carried out by a proper fusion of measurements from different sensors (i.e., accelerometers, gyroscopes, and magnetometers). The performance of inertial systems is good when all types of sensors are used but the cost tends to increase when using a large number of sensors.[12] However, if one attempts to reduce system costs (related to the use of multiple sensors per device) by using only partial subsets of the previous sensors, the performance of inertial systems decreases because of implicit limitations in determining a stable and complete orientation of the devices when using independently each considered sensor. For instance: accelerometers cannot provide information about the devices' heading; magnetometers suffer from magnetic fields disturbance; and gyroscopes present a bias drift.

The idea behind our work is that of reducing the cost (and complexity) of standard inertial systems by designing a BAN-based system where devices are equipped with only an accelerometer (and a radio interface). To this end, note that the proposed system does not require computationally intensive operations. Indeed, the operations involved in the algorithms can be executed at most in polynomial time complexity (mostly due to the trigonometric operations on the inertial measurements). Note also that the proposed system "as-is," when deployed on SunSPOTs, is already able to run with a time delay of few tens of milliseconds (i.e., almost real-time), even with a not completely optimized code implementation.

The proposed system would be suitable to estimate the arm inclination but cannot provide information about its heading. To this end, radio localization is also used in order to localize (upon proper training of the system) the devices and, then, to estimate the corresponding arm segment heading. The proposed hybrid approach allows to inherit the advantages of inertial systems (in terms of robustness against typical limitations of optical systems) but, on the other hand, introduces a major limitation. Indeed, due to the fact that radio localization relies on a preliminary training phase and anchor nodes are fixed in the surroundings, the user is forced to keep his/her body (with the exception

[12] Note, for instance, that the cost of the majority of commercial devices used for motion capture purposes (i.e., typically equipped with an accelerometer, a gyroscope, and a magnetometer) is at least twice (or even three times) that of a simpler device equipped just with an accelerometer. In fact, the cost of a gyroscope (or of a magnetometer) alone is already twice that of an accelerometer.

of the arm) still during the arm posture recognition session (or, at least, he/she is forced to keep the shoulder, corresponding to the system origin, still).

Even if the requirement of still users would not be a true limitation for several rehabilitation applications (where users are not supposed to move around when performing specific exercises), some applications may require that the user moves while performing an exercise. Therefore, an appealing extension of the proposed system should aim at allowing the user to move freely while performing the exercises. As a first step in this direction, in this paper we have also considered anchor nodes placed directly on the user body and verified that they can be effectively used as reference nodes. Furthermore, a future design of the proposed system could exploit radio localization to directly estimate distances between (mobile) nodes placed only on the user body, possibly removing the distinction between anchor and target nodes (i.e., every node could be at the same time an anchor and a target node). This could have a direct relevance to rehabilitation engineering and is currently under investigation.

Another possible limitation of the proposed system, especially if the application scenario involves evaluating elderly people movements, is the required training phase, which could be sometimes hard to be correctly performed by the user. To this end, a possible solution may consist in allowing an impaired user to perform it and then rely on techniques able to "map" the training output in order to be used with the actual patient (similar techniques have already been taken into account for Kinect). In general, since this work represents a feasibility study on the hybrid integration of radio and inertial signals, the system "as-is" is not intended to be directly applied in real scenarios (e.g., rehabilitation applications) and, therefore, further extensions and methods should be considered in order to improve its robustness.

Our results show that the proposed system, in its current design, can provide a performance similar to that of Kinect (which is a concurrent low-cost posture recognition system). In addition, simple localization algorithms, namely deterministic (the kNN algorithm) and probabilistic (p-kNN), have been considered to recognize arm postures – due to their low computational complexity, these algorithms can be implemented on the majority of current low-cost devices. The system parameters have been optimized in order to minimize the average PE and values around 10 cm have been obtained. A PE around 10 cm can surely be sufficient for the majority of posture *classification* applications. The system performance has also been evaluated in terms of precision and accuracy. In particular, the deterministic version of the proposed hybrid localization algorithm outperforms the probabilistic one and slightly outperforms an optical Kinect system.

Since the aim of posture *classification* applications is just to determine the posture of a user choosing among a discrete set of previously trained postures (and, thus, to discriminate between known postures), our system works properly if the considered postures are sufficiently spatially distinct (i.e., the distance, for each pair of different postures, of at least two corresponding body segments equipped with sensor nodes is higher than 10 cm). Concerning posture *recognition* (where one is not interested on the discrimination between priorly known discrete and well-distinguished postures but, rather, on the recognition of generic postures), the performance of our system highly depends on the application requirements. In particular, it could be used to monitor the recovery improvement after an arm surgery, in the cases where it is relevant to discriminate a few cases (e.g., the arm can be raised half-way or all the way). For applications where it is necessary to discriminate between very close positions, then the accuracy needs to be improved. To this end, possible ways to improve the system performance may reside, for instance, in the use of an outlier rejection technique to handle misleading RSS measurements and/or in the use of advanced filtering techniques (which should give adaptive weights to every different measurement) in order to properly fuse together inertial and radio signals [28].

References

[1] G.M. Djuknic and R.E. Richton, Geolocation and assisted GPS, *IEEE Computer* **34**(2) (February 2001), 123–125.

[2] P. Bahl and V.N. Padmanabhan, RADAR: An in-building RF-based user location and tracking system, in: *Proc. IEEE Conf. on Computer Commun. (INFOCOM)*, Tel Aviv, Israel, Vol. 2, March 2000, pp. 775–784.

[3] K. Kaemarungsi, Design of indoor positioning systems based on location fingerprinting technique, Ph.D. dissertation, University of Pittsburgh, Pittsburgh, 2005, available at http://etd.library.pitt.edu/ETD/available/etd-02232005-235903/unrestricted/dissertation28Feb05.pdf.

[4] N. Swangmuang and P. Krishnamurthy, Location fingerprint analyses toward efficient indoor positioning, in: *IEEE Int. Conf. Pervasive Comp. and Commun. (PERCOM)*, Hong Kong, China, March 2008, pp. 100–109.

[5] M. Giuberti, M. Martalò and G. Ferrari, Fingerprinting-based wireless 3D localization for motion capture applications, in: *Proc. 1st ACM MobiHoc Workshop on Pervasive Wireless*

Healthcare (MobileHealth), Paris, France, May 2011, pp. 6.1–6.8.
[6] M. Martalò, M. Giuberti and G. Ferrari, Experimental investigation of wireless sensor networks for fingerprinting-based posture recognition, in: *Riunione Annuale 2011 del Gruppo Nazionale Telecomunicazioni e Teoria Dell'Informazione (GTTI)*, Messina and Taormina, Italy, June 2011.
[7] S.O.H. Madgwick, An efficient orientation filter for inertial and inertial/magnetic sensor arrays, Department of Mechanical Engineering, University of Bristol, Bristol, UK, Tech. rep., April 2010, available: http://www.x-io.co.uk/res/doc/madgwick_internal_report.pdf.
[8] H.J. Luinge, P.H. Veltink and C.T.M. Baten, Ambulatory measurement of arm orientation, *Journal of Biomechanics* **40**(1) (January 2007), 78–85.
[9] D. Roetenberg, P.J. Slycke and P.H. Veltink, Ambulatory position and orientation tracking fusing magnetic and inertial sensing, *IEEE Trans. on Biomedical Eng.* **54**(5) (May 2007), 883–890.
[10] Motion capture systems from Vicon, http://www.vicon.com/.
[11] Microsoft Kinect, http://www.xbox.com/en-US/kinect/.
[12] C. Chang, B. Lange, M. Zhang, S. Koenig, P. Requejo, N. Somboon, A. Sawchuk and A.A. Rizzo, Towards pervasive physical rehabilitation using Microsoft Kinect, in: *Proc. of the 6th Int. Conf. on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, San Diego, CA, USA, May 2012, pp. 159–162.
[13] T.B. Moeslund, A. Hilton and V. Krüger, A survey of advances in vision-based human motion capture and analysis, *Computer Vision and Image Understanding* **104**(2) (2006), 90–126.
[14] L. Chen, H. Wei and J. Ferryman, A survey of human motion analysis using depth imagery, *Pattern Recognition Letters* (2013), doi:10.1016/j.patrec.2013.02.006.
[15] D. Roetenberg, H.J. Luinge and P.J. Slycke, Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors, *Xsens Technologies B.V.* (April 2009).
[16] D.A. Winter, *Biomechanics and Motor Control of Human Movement*, John Wiley & Sons, New York, NY, USA, 1990.
[17] M. Brodie, A. Walmsley and W. Page, Fusion motion capture: A prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing, *Sports Technology* **1**(1) (June 2008), 17–28.
[18] Y. Tao, H. Hu and H. Zhou, Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation, *The International Journal of Robotics Research* **26**(6) (June 2007), 607–624.
[19] D. Vlasic, R. Adelsberger, G. Vannucci, J. Barnwell, M. Gross, W. Matusik and J. Popović, Practical motion capture in everyday surroundings, *ACM Trans. on Graphics* **26**(3) (July 2007), 35.1–35.9.
[20] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas and J. Sievanen, A probabilistic approach to WLAN user location estimation, *Int. Journal Wireless Information Networks* **9**(3) (July 2002), 155–164.
[21] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification and Scene Analysis*, 2nd edn, Wiley-Interscience, New York, NY, USA, 2000.
[22] D. Shepard, A two-dimensional interpolation function for irregularly spaced data, in: *ACM National Conf.*, 1968.
[23] J.B. Kuipers, *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality*, Princeton University Press, Princeton, NJ, USA, 1999.
[24] X. Yun, E.R. Bachmann and R.B. McGhee, A simplified quaternion-based algorithm for orientation estimation from earth gravity and magnetic field measurements, *IEEE Trans. on Instrumentation and Measurement* **57**(3) (March 2008), 638–650.
[25] SunSPOT World, http://www.sunspotworld.com.
[26] J. Hightower and G. Borriello, Location systems for ubiquitous computing, *IEEE Computer* **34**(8) (August 2001), 57–66.
[27] Š. Obdržálek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison and M. Pavel, Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population, in: *Proc. of the 34th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society (EMBS)*, San Diego, CA, USA, August 2012, pp. 1188–1193.
[28] F. Gustafsson, *Statistical Sensor Fusion*, Studentlitteratur, Lund, Sweden, 2010.