Information Fusion 41 (2018) 243–254

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Geometrical kinematic modeling on human motion using method of multi-sensor fusion

Cheng Xu[a,b,1], Jie He[a,b,1,*], Xiaotong Zhang[a,b,*], Cui Yao[a,b], Po-Hsuan Tseng[c]

[a] School of Computer and Communication Engineering, University of Science and Technology Beijing, China
[b] Beijing Key Laboratory of Knowledge Engineering for Materials Science, China
[c] Department of Electronic Engineering, National Taipei University of Technology, China

ARTICLE INFO

Article history:
Received 14 January 2017
Revised 11 September 2017
Accepted 12 September 2017
Available online 14 September 2017

Keywords:
Human motion
Geometrical characteristics
Denavit–Hartenberg convention
PCRLB
Wearable sensors

ABSTRACT

Human motion sensing based on wearable sensors could be viewed as a multi-objects tracking issue of human body joints. Sensor drift errors and distortion are the main challenges of the tracking accuracy of human motion. Traditional filtering and fusion methods, such as Kalman filter, can to some extent reduce the instantaneous error but cannot avoid sensor drift fundamentally. Physical characteristics of human body should be considered and human motion models should be exhibited to describe human motions. The existing models such as skeleton model and cylinder model are either too simple or too complicated for practical applications. In this study, we put forward a geometrical kinematic characteristics based human motion model. The whole human body is viewed as an articulated skeleton and Denavit–Hartenberg convention is adopted to describe the forward kinematics structure. Theoretical analysis is conducted with the derivation of Posterior Cramer–Rao Lower Bound (PCRLB) in human movement scenes based on proposed model. Significant superiority is shown in simulation results. An experiment on human lower limb motions is carried out to verify the validity of the proposed human motion model in practice applications, from the angles of both capturing accuracy and energy consumption. The capturing accuracy has an obvious increase in the testing results, with acceptable energy consumption. It is far more efficient than traditional methods.

1. Introduction

With the booming of body sensor networks (BSNs), people-centric motion sensing applications have been developing very fast. Due to the advantages of low cost and small size, wearable sensors have been drawing attention. They can meet most needs of human motion monitoring and have been adopted in many applications, such as interactive gaming and learning, animation, health care, personal navigation and security monitoring [1–9].

As BSN systems can be directly used to monitor several vital signs continuously and noninvasively, with sensors placed on body parts, they are widely used in e-Health applications [2–4]. These signals can be captured by multiple sensors, and in turn, infer many diseases at an early stage. Sameer et al. [4] defined a framework that managed common tasks for healthcare monitoring applications to aid development of BSN. Chen et al. [5] proposed body temperature sensor that was wearable, breathable and stretchable for healthcare monitoring. They are capable of measuring temperature, pressure and strain so far. Bebek et al. [6] and Bamberg et al. [7] utilized a pressure sensor array and inertial measurement unit (IMU) for gait analysis. Ghasemzadeh et al. [8] introduced a novel classification model that identified physical movements from body-worn inertial sensors while taking collaborative nature and limited resources of the system into consideration. Inertial sensors are integrated with ultrawideband (UWB) localization system in [9] for simultaneous 3-D trajectory tracking and lower body motion capture (MoCap) under various dynamic activities such as walking and jumping. This method obtained accurate results but also introduced additional complexity.

Despite the advantages mentioned above, there still exist challenges when wearable sensors are applied to human motion sensing. As human motion sensing could be viewed as a multiple targets localization issue of human body joints, tracking accuracy is the most important consideration. However, sensor drift errors and distortion (especially in long time monitoring) are the main problems.

* Corresponding authors at: School of Computer and Communication Engineering, University of Science and Technology Beijing, China.
E-mail addresses: hejie@ustb.edu.cn (J. He), zxt@ustb.edu.cn (X. Zhang).
[1] Cheng Xu and Jie He are co-first authors and contributed equally to this work.

http://dx.doi.org/10.1016/j.inffus.2017.09.014
1566-2535/

Kalman filter and calibration algorithms are the most widely adopted method to overcome drift errors. They are both used to overcome the instantaneous error problem and multiple sensor fusion problem. Luinge et al. [10] realized a method to measure the orientation of human body segments using Kalman filter that took into account the spectra of the signals, as well as a fluctuating gyroscope offset, and thereby improved the estimation accuracy. However, the reduction of instantaneous errors delayed the accumulation but cannot avoid the sensor drift fundamentally. Since human motion sensing could be viewed as multiple targets tracking issue, fusion methods [11–13] are studied to fuse multiple channel information. However, these studies to some degree improve the recognition accuracy in the angle of intelligent algorithms, but cannot meet the drawbacks of data source. Errors still accumulate as time goes on [11].

To better solve these problems, expedient models should be exhibited to describe human motions. Both filter and intelligent recognition algorithms rely on system models [10]. The expecting models should fulfill the following requirements:

- The description method of the motion model should be adaptable to different persons, namely with good generalization capability.
- It's possible to incorporate prior known constraints to narrow down the search space of algorithms and improve the capturing accuracy.
- The entire human body should be viewed as a whole and the correlation connected body relationship should be considered.

Based on above considerations, skeleton model is widely adopted in human motion monitoring studies and applications, especially in computer vision area [1,10]. The human body can be modeled as an articulated structure with 15–19 limbs [14,15], and the overall posture of the human subject can be accurately acquired by determining the orientation of each limb. However, in these studies, motion recognition systems rely on fine-grained tracking on exactly human joints using cameras [1] and inertial sensors [10]. Tracking is always conducted solely on human joints or body parts. Commercial systems like Xsens [16] and InterSense [17] employ similar approaches in commercial applications with monitoring each sole joint's movements using inertial and magnetic sensors or ultrasonic sensors. Cylinder model is another common adopted human motion model, quoted in literatures [18] and [19]. Each body part is represented with a degenerated cylinder and the joints are described with various degrees of freedoms(DoFs). The overall body model is built in a tree-like hierarchy starting with the torso as root body part. Each child is described with a degenerated cylinder and the corresponding transformation from its parent. Joint model is also introduced in [20] and ICP-based approach is applied for pose estimation. However, it seems that the effect of the algorithm is partially removed when the constraints are enforced. In addition, the multiple DoFs information could hardly be obtained in practical applications.

Considering the above problems, since skeleton model takes body joints solely and cylinder model do not work well when constraints are applied, improvements should be developed. The proposal of a structured human motion model is urgent, considering both geometrical characteristics and the integrality of human body. A kinematics-based model is a better choice. Correspondingly, body connections are considered in [21,22]. Mihelj et al. [21] considered a technique for computation of the inverse kinematic model of the human arm. The algorithm enables estimation of human arm posture, but the drift errors are never considered. Vlasic et al. [22] presented a system for acquiring motions with wearable sensors gathering ultrasonic time-of-flight and inertial measurements. The information is combined using an Extended Kalman Filter to reconstruct joint configurations of a body. However, the ultrasonic sensors frequently suffer from serious signal interference.

With the requirements for motion sensing in mind, we take the advantages of both skeleton and cylinder models, as well as considering kinematic characteristics of human body. In this study, we propose a geometrical characteristics based articulated human model, considering the co-movements of various parts of the body. The main contribution of this paper is using the geometrical constrained method (GC method) and the connectivity between joints to improve human motion monitoring accuracy. The basic idea is to use Denavit–Hartenberg Convention forward kinematics to reconstruct human body and revise the estimation error and distortion. D-H transformation makes use of the relative and absolute positions between joints and limbs and these linkage constraints to some extent limit the drift errors of sensors.

The rest of this paper is organized as follows. In Section 2, we analyzed human motions kinematic characteristics, and proposed the geometrical model, as well as its representation. In Section 3, PCRLB simulation is presented, and verify the theoretical performance of proposed model. Field experiment and results analysis are given in Section 4 for human lower limb analysis and personal navigation. Finally, conclusion comes in Section 5.

2. Modeling of human motions

As mentioned above, the whole human body is an assembly of body parts connected to manage forces and movements. Each motion is accomplished collaboratively rather than independently. The movement of a body is studied using geometry so the limb is considered to be rigid. The connections between limbs are called joints. A linkage modeled as a network of rigid limbs and ideal joints is called a kinematic chain. In order to understand better the human motion, in this chapter, we take a deep analysis on geometrical characteristics of human body. We model the human body as a combination of joints and limbs, and the coordinate description method is proposed. With the proposed kinematic model, geometrical information is merged into consideration and the human body is more easily to be viewed as a whole.

2.1. Geometrical kinematic body description

Generally, a 3D human motion is considered as a series of continuous movements of joints and limbs. These motions are mathematically presented by coordinates and orientations of sequential body parts. The whole body could be regarded as an articulated chain structure composed of several rigid segments(limbs) and joints [24,25]. This kind of chain structure makes the movement of each part of the body not independently but cooperatively. The articulated human body has a large number of kinematic joints, but a particular motion may only involve a small subset of them. The possible related body parts could be represented as a linear combination of the joints. Therefore, in order to clearly describe human motions, we group the estimated body into five parts (left-arm, right-arm, trunk, left-leg, right-leg) and signify human bodies with dots and lines. The five parts together contribute to a pyramid-like whole body. Detailed description is shown in Fig. 1.

The rigid body skeleton is modeled as a tree structure, consisting of five connected parts. Each part is generally consist of three joints and two limbs. Then we define a set of joints and lines on the human skeleton. There are 15 joints and 14 limbs.

Joints. Each joint J is encoded with its coordinate(x_J, y_J, z_J).

Limbs. These lines $L_{J_1 \to J_2}$ refer to virtual space position between joints J_1 and J_2. One of these two constraints should be satisfied:

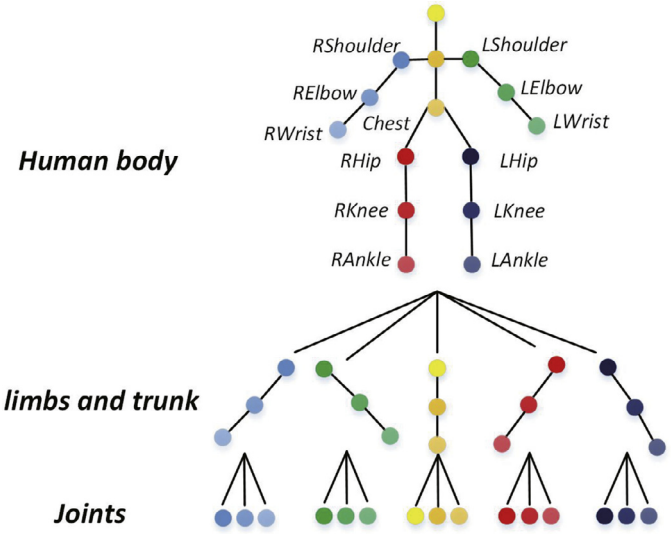


Fig. 1. Kinematic articulated human skeleton model. A complete human body is described from three levels: 1) Joints. 2) Limbs and trunk. 3) Whole body.

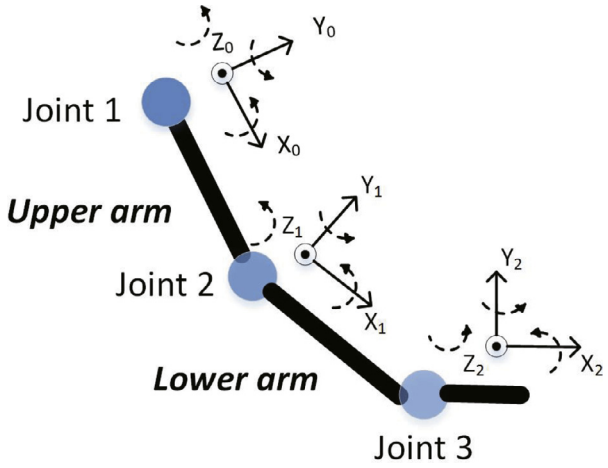


Fig. 2. Human motion coordinate system. Take Arm as an example. Local motion parameters are mainly join coordinates, angles and relative distance.

1) J_1 and J_2 are directly adjacent in the kinetic chain. These limbs are clearly drawn in Fig. 1 with black solid line.
2) One of J_1 and J_2 is two or more steps away on the same kinetic chain (i.e., $L_{LShoulder \rightarrow LWrist}$ and $L_{RHip \rightarrow RAnkle}$). These lines significantly reflect the relative relationship between indirect connected joints.

In the description above, we mainly expound how human body could be represented with basic elements, namely joints and limbs. A linkage structure is constructed and every human motion could be represented as a cooperative movement of connected joints and limbs. The movement of an ideal joint is generally associated with a subgroup of Euclidean displacements. Two or more rigid bodies in space are collectively called a rigid body system. Among numerous complex motions, arms and legs are the most important participants and they conduct lots of actions. For example, legs(or lower limb) play a big role in motions of walking, jogging, running and jumping, whereas arms(or upper limb) matter more in motions of swing hands and lifting. Moreover, motions like climbing up and down are in need of the participation of both lower and upper limbs. However, in majority the trunk remains relatively steady in compare to the end joints or limbs of the human body.

This phenomenon is quite similar with the end-effect problem [29] in robotics area. The action of humanoid robots is usually reconstructed based on kinematic constraints and the forward kinematic method. Similarly, on the issue of human motion sensing, we can describe the motion of these independent rigid bodies with kinematic constraints. Kinematic constraints are relations between rigid bodies, namely the geometrical coordinates (x_J, y_J, z_J) and distance constraints $L_{J_1 \rightarrow J_2}$. The Denavit–Hartenberg convention parameters (also called D–H parameters) are introduced to describe kinematic human motion. It is widely used in humanoid robot-motion reconstruction and articulated mechanical motion monitoring. In the following part of this section, the detail definition of coordinate system and D–H parameters will be demonstrated.

2.2. Coordinate system definition

A local coordinate system is the most important description component as it meticulously describes the details of human actions. Under the local system, in order to show the relative position between each joint, on the basis of our proposed segmented skeleton model, the spine (chest) is selected as the origin of the whole local system. As the segmented skeleton has a layered structure, this origin is equivalently the root node of a tree. Meanwhile, since the movement of each joint has its independence, independent sub-coordinate system originating as each joint can be constructed in direction of its movement. At the same time, because the various joints are linked together by bones, their respective movement cannot be totally separated but have certain associations. This conforms to the mechanical system of connecting rod movement, using Denavit–Hartenberg dynamics equation. Since then, we have started from a kinetic analysis and model human motions with mathematical representations.

In the following part of this chapter, we will demonstrate how local coordinate system is constructed and what parameters from this system we could obtain to describe human motions.

2.2.1. Coordinate system construction

Fig. 2 is shown as an example. Local coordinate system should be constructed following these principles:

1. Number the joints from 1 to n starting with the base and ending with the end-effector.
2. ***Establish the base coordinate system.*** Establish a right-handed orthonormal coordinate system (X_0, Y_0, Z_0) at the supporting base with Z_0 axis lying along the axis of motion of joint 1.
3. ***Establish joint axis.*** Align the Z_i with the axis of motion (rotary or sliding) of joint $i+1$.
4. ***Establish the origin of the ith coordinate system.*** Locate the origin of the ith coordinate at the intersection of the Z_i & Z_{i-1} or at the intersection of common normal between the Z_i & Z_{i-1} axes and the Z_i axis.
5. ***Establish $X_i$axis.*** The X_i axis lies along the common normal from the Z_{i-1} axis to the Z_i axis, $X_i = \pm(Z_{i-1} \times Z_i)/\|Z_{i-1} \times Z_i\|$), (if Z_{i-1} is parallel to Z_i, then X_i is specified arbitrarily, subject only to X_i being perpendicular to Z_i);
6. ***Establish $Y_i$axis.*** Assign $Y_i = +(Z_i \times X_i)/\| Z_i \times X_i \|$) to complete the right-handed coordinate system.
7. The hand coordinate frame is specified by the geometry of the end-effector. Normally, establish Z_n along the direction of Z_{n-1} axis and pointing away from the robot; establish X_n such that it is normal to both Z_{n-1} and Z_n axes. Assign Y_n to complete the right-handed coordinate system.

Local coordinate system could be constructed based upon above procedures. Then motion parameters generated in this system is described as follows.

2.2.2. Parameters definition

The following four transformation parameters are used to Limb and Joint motions, which are known as Denavit–Hartenberg(D–H) parameters [26]:

1. Joint angle θ_i: the angle of rotation from the X_{i-1} axis to the X_i axis about the Z_{i-1} axis. It is the joint variable if joint i is rotary.
2. Joint distance d_i: the distance from the origin of the $(i-1)$ coordinate system to the intersection of the Z_{i-1} axis and the X_i axis along the Z_{i-1} axis. It is the joint variable if joint i is prismatic.
3. Limb length a_i: the distance from the intersection of the Z_{i-1} axis and the X_i axis to the origin of the ith coordinate system along the X_i axis.
4. Limb twist angle α_i: the angle of rotation from the Z_{i-1} axis to the Z_i axis about the X_i axis.

With these parameters, human body parts are possible to be viewed as a geometrical connected whole. Comovement relation between joints and limbs is covered and constrained because relative information is taken into consideration. How these effects work will be detailed in the next chapter.

Once the process model and measurement model are determined, the rest of the problem is to consider how to obtain the four D–H parameters. The parameters, θ and α, are easily obtained with the fusion of IMU sensors, namely accelerometer and gyroscope [8]. The parameters, d and a, can be achieved with TOA(time of flight) ranging method. Ultra-wide Bandwidth(UWB) is widely used in high accuracy distance measuring applications [9]. In the following study, IMU sensors and UWB ranging nodes are fused to capture the D-H parameters needed by proposed motion model.

2.3. Geometrical kinematic model

Based on above discussion, let $m_k = [P, n]^T = [{}^X P_k, {}^Y P_k, {}^Z P_k, {}^X n_k, {}^Y n_k, {}^Z n_k]^T$, $(k = 1, \ldots, K)$ be the state vector at the kth state, where $({}^X P_k, {}^Y P_k, {}^Z P_k)$ is the 3D location coordinate of the joint in world coordinate space and $({}^X n_k, {}^Y n_k, {}^Z n_k)$ is the norm vector indicating direction of joint movement, for short of matrix R. K is the total number of observation frames. We define $\overline{\hat{m}}_k$, as the predicted position of the joint from sensor based joint tracking and define $\hat{z_k}$ as the measured position from RF ranging based localization. Thus, the system state transition function can be given as

$$\begin{cases} \overline{\hat{P}}_{k+1} = A_k \hat{m}_k + q_k \\ z_{k+1} = H_k \overline{\hat{m}}_k + r_k \end{cases} \tag{1}$$

where q_k is the inaccuracy of movement estimation that follows Gaussian distribution with covariance Q and r_k is the inaccuracy of RF based location estimation that follows Gaussian distribution with covariance R. The matrix A_k and matrix H_k can be given as

$$A_k = \begin{bmatrix} 1 & 0 & 0 & \hat{d}_k & 0 & 0 \\ 0 & 1 & 0 & 0 & \hat{d}_k & 0 \\ 0 & 0 & 1 & 0 & 0 & \hat{d}_k \\ 0 & 0 & 0 & & & \\ 0 & 0 & 0 & & [R] & \\ 0 & 0 & 0 & & & \end{bmatrix}^T$$

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T$$

where $R = RR_k R^{-1}$ is the accumulative rotation matrix and R_k is the standard rotation matrix that contains yaw–pitch–roll information.

With the use of parameters in defined coordinate systems, relative positions of different human motions can be represented. The denoted information consists of two parts, the Rotation vector $\mathbb{R}$ and the Position vector $\mathbb{P}$, namely

$$M_0^n = M_0^1 M_1^2 M_{n-1}^n = \begin{bmatrix} R_0^n & P_0^n \\ 0 & 1 \end{bmatrix} \tag{2}$$

where rotation vector is defined as a 3*3 matrix, $\mathbb{R} = \begin{bmatrix} \boldsymbol{n} & \boldsymbol{o} & \boldsymbol{a} \end{bmatrix} = \begin{bmatrix} n_x & o_x & a_x \\ n_y & o_y & a_z \\ n_z & o_z & a_z \end{bmatrix}$. Position vector $\mathbb{P}$ directs to the heart of local coordinate system from the origin of reference system, namely $P = \begin{bmatrix} \boldsymbol{p_x} & \boldsymbol{p_y} & \boldsymbol{p_z} \end{bmatrix}^T$.

Since we defined joint parameters such as Joint angle θ_i and Joint distance d_i, time-variant characteristics of human motions should be represented by these data. Four successive elementary transformations are required to relate the ith coordinate frame to the $(i-1)$th coordinate frame:

1. Rotate about the Z_{i-1} axis an angle of θ_i to align the X_{i-1} axis with the X_i axis, namely $R(z_{i-1,\theta_i})$.
2. Translate along the Z_{i-1} axis a distance of d_i, to bring X_{i-1} and X_i axes into coincidence, namely $T(z_{i-1}, d_i)$.
3. Translate along the X_i axis a distance of a_i to bring the two origins O_{i-1} and O_i as well as the X axis into coincidence, namely $R(x_i, \alpha_i)$.
4. Rotate about the X_i axis an angle of α_i (in the right-handed sense), to bring the two coordinates into coincidence, namely $T(x_i, a_i)$.

The rotation matrix between joint i and joint $i-1$ can be written as:

$$\begin{aligned} M_{i-1}^i &= \begin{bmatrix} R_{i-1}^i & P_{i-1}^i \\ 0 & 1 \end{bmatrix} \\ &= R(z_{i-1}, \theta_i) T(z_{i-1}, d_i) R(x_i, \alpha_i) T(x_i, a_i) \\ &= \begin{bmatrix} \cos\theta_i & -\cos\alpha_i \sin\theta_i & \sin\alpha_i \sin\theta_i & a_i \cos\theta_i \\ \sin\theta_i & \cos\alpha_i \cos\theta_i & -\sin\alpha_i \cos\theta_i & a_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{3}$$

We denote a 3D position vector to represent human motion as $P = \begin{bmatrix} X & Y & Z \end{bmatrix}^T$. Furthermore, as mentioned before, a time-domain motion could be viewed as a series of transformation from a moment before. As the position transformation from joint $k-1$ to joint k may be written as

$$P_k = P_{k-1}^k + R_{k-1}^k P_{k-1} \tag{4}$$

Based on former definition, position at joint k could be further represented as

$$\begin{bmatrix} {}^X P_k \\ {}^Y P_k \\ {}^Z P_k \end{bmatrix} = \begin{bmatrix} {}^X P_{k-1}^k \\ {}^Y P_{k-1}^k \\ {}^Z P_{k-1}^k \end{bmatrix} + \begin{bmatrix} n_x & o_x & a_x \\ n_y & o_y & a_z \\ n_z & o_z & a_z \end{bmatrix} \begin{bmatrix} {}^X P_{k-1} \\ {}^Y P_{k-1} \\ {}^Z P_{k-1} \end{bmatrix} \tag{5}$$

Namely, it could be viewed as the result of Translation and Rotation from the position at joint $k-1$. To simplify the set up, its defined as

$$\begin{aligned} \begin{bmatrix} {}^X P_k \\ {}^Y P_k \\ {}^Z P_k \\ 1 \end{bmatrix} &= \begin{bmatrix} n_x & o_x & a_x & {}^X P_{k-1}^k \\ n_y & o_y & a_z & {}^Y P_{k-1}^k \\ n_z & o_z & a_z & {}^Z P_{k-1}^k \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^X P_{k-1} \\ {}^Y P_{k-1} \\ {}^Z P_{k-1} \\ 1 \end{bmatrix} \\ &= M_{k-1}^k \begin{bmatrix} {}^X P_{k-1} \\ {}^Y P_{k-1} \\ {}^Z P_{k-1} \\ 1 \end{bmatrix} \end{aligned} \tag{6}$$

Further simplification is possible. The position at joint k could be

a series of transformation from the original joint numbered 0 after Translation and Rotation. So a structured human motion could be finally modelled as

$$\begin{bmatrix} {}^{X}P_k \\ {}^{Y}P_k \\ {}^{Z}P_k \\ 1 \end{bmatrix} = M_0^k \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

If we define the priori estimate error covariance as $\hat{\mathcal{E}} = E[(P_k - \overline{\hat{m}}_k)(P_k - \overline{\hat{m}}_k)^T]$ and define the posteriori estimate error covariance as $\mathcal{E} = E[(P_k - \overline{m}_k)(P_k - \overline{m}_k)^T]$, a classic Kalman filter [27] can be exploited to hybrid the information from both sensor sensing based and RF based localization. The mathematical description of the hybrid motion tracking approach is shown in Algorithm 1 and

Algorithm 1 Kalman filtering for D–H based motion tracking.

Input: : Initialized $\hat{z}_1$ and P_1, $\hat{m}_1 = \hat{z}_1$.
Output: : Expected $\hat{m}_K$.
1: **for** $k = 2$ to K **do**
2: Predict the state, $\overline{\hat{m}}_k = A\hat{m}_k$
3: Predict state error covariance, $\hat{\mathcal{E}}_k = A\mathcal{E}_{k-1}A^T + Q$
4: Update the rotation matrix, $R = RR_kR^{-1}$
5: Compute Kalman gain, $K_k = \hat{\mathcal{E}}_k H^T (H\hat{\mathcal{E}}_k H^T + M)^{-1}$
6: Update the state, $\hat{m}_k = \overline{\hat{m}}_k + K_k(z_k - H\overline{\hat{m}}_k)$
7: Update state error covariance, $\mathcal{E}_k = (I - K_kH)\overline{\mathcal{E}}$
8: **end for**

the performance of the hybrid approach will be discussed later in the next Section.

3. PCRLB simulation

The motion capture accuracy is the most important metric to evaluate the proposed model's performance. Human motion monitoring is a kind of tracking on the human body joints and limbs. It is reasonable to view it as a localization problem. Posterior Cramer–Rao Lower Bound (PCRLB) [33] is one of the most widely used criteria to evaluate models and algorithms [23,28]. It offers the minimum limit estimation error of all unbiased estimators and can perfectly show the performance of models and algorithms. However, to the best of our knowledge, in body sensor network (BSN) areas, no related studies have been done on the theoretical analysis of the error bounds of motion capture, especially when distance and inertial measurement are both taken into consideration. In this section, we first analyze the errors in motion capture, then derive the PCRLB of proposed human motion model. Simulations of performance comparison are performed to verify the effectiveness of this model.

For the sake of convenience , in the following derivation process, three 3D space parameters are introduced in representation of the four D–H convention parameters. The parameters, respectively, are distance measurement d, azimuth angle θ and pitch angle α. The definition of these parameters is shown in Fig. 3. we define $\nabla_a \triangleq [\frac{\partial}{\partial a_1}, \ldots, \frac{\partial}{\partial a_M}]^T$, $\nabla_a^b \triangleq \nabla_b \nabla_a^T$. $p(a)$ the probability density function (p.d.f.) of the random vector a, $\otimes$ the Kronecker product, $Tr\{\cdot\}$ the matrix trace.

3.1. Error definition

Human motion capture is the monitoring of various joints along the time sequence. The tracking position of joint at time k is denoted as P_k. In time line, the tracking of joint k could be described as

$$P_{k+1} = P_k + d_kT_k + s_k \tag{8}$$

where s_k is additive Gaussian noise with the mean μ_0 and standard deviation σ_0, namely $s_k \sim \mathcal{N}(\mu_0, \sigma_0^2)$.

The coefficient vector T_k is $T_k = [\sin\alpha_k\cos\theta_k, \sin\alpha_k\sin\theta_k, \sin\phi_k]^T$ where the azimuth θ_{k-1} and the pitch ϕ_{k-1} can be calculated by IMU sensors using gyroscopes and magnetometers.

$$\hat{d}_k = d_k + n_k \tag{9}$$

where d_k is the actual Euclidean distance. The distance measurement noise comply with Gaussian distribution, namely $n_k \sim \mathcal{N}(\mu_d, \sigma_d^2)$. n_k is the measurement noises that can be modeled as uncorrelated zero-mean Gaussian random variable with variances σ_d^2. Vector $\hat{\boldsymbol{d}} = [\hat{d}_0, \hat{d}_1, \ldots, \hat{d}_{k-2}]^T$ is introduced to collect $\hat{d}_k$.

Horizontal heading estimates from inertial sensor based approach reads

$$\hat{\theta}_k = \theta_k + u_k, u_k \sim N(0, \epsilon_k^2) \tag{10}$$

where θ_k is the actual horizontal heading. u_k is uncorrelated zero-mean Gaussian random variable with variances ϵ_k^2, which is independent from z direction and is also uncorrelated with $\hat{d}$. Vector $\hat{\theta}_k[\hat{\theta}_0, \hat{\theta}_1, \ldots, \hat{\theta}_{k-2}]^T$ is introduced to collect $\hat{\theta}_k$.

Vertical elevation estimates from inertial sensor based approach reads

$$\hat{\alpha}_k = \alpha_k + v_k, v_k \sim N(0, \xi_k^2) \tag{11}$$

where α_k is the actual vertical elevation. v_k is uncorrelated zero-mean Gaussian random variable with variances ξ_k^2, which is independent from z direction and is also uncorrelated with both $\hat{\theta}$ and $\hat{d}$. Vector $\hat{\alpha}_k[\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_{k-2}]^T$ is introduced to collect $\hat{\alpha}_k$.

3.2. Posterior Cramer–Rao lower bound

As is mentioned at the beginning of this section, regarding the time-variant statistical system models for hybrid motion capturing, CRLB is most frequently used to provide the performance lower bound. With posterior information of random parameters such as step length, heading and elevation, CRLB can be extended to PCRLB. In this section, we introduce the mathematical basis for PCRLB calculation.

The best place to start PCRLB derivation is the joint p.d.f. as follows:

$$\begin{aligned} p(\hat{r}, \hat{d}, \hat{\theta}, \hat{\alpha}, \hat{P}) = p(\hat{r}_0|P_0) \prod_{k=0}^{K-2} p(\hat{d}_k|P_{k+1}, P_k) \\ p(\hat{\theta}_k|P_{k+1}, P_k)\, p(\hat{\alpha}_k|P_{k+1}, P_k)\, p(\hat{r}_{k+1}|P_{k+1}) \end{aligned} \tag{12}$$

for which the PCRLB of estimating joint location P can be given as

$$E([\hat{P}(u) - P][\hat{P}(u) - P]^T) \geq F(P)^{-1} \tag{13}$$

where $\hat{P}$ denotes to estimated joint location and P denotes the actual joint location as given in [(13)]. $F(P)$ is the Fisher information matrix (FIM) given as

$$F = E_{\hat{r},\hat{d},\hat{\theta},\hat{\alpha}}[-\nabla_P^P ln p(\hat{r}, \hat{d}, \hat{\theta}, \hat{\alpha}, \hat{P})] \tag{14}$$

To calculate the FIM at kth step $F(P_{0:\,k})$, we define $p_k = p(\hat{r}_{0:k}, \hat{d}_{0:k}, \hat{\theta}_{0:k}, \hat{\alpha}_{0:k}, \hat{P}_{0:k})$, where $\hat{r}_{0:k} \triangleq [\hat{r}_0, \hat{r}_1, \ldots, \hat{r}_k]^T$, $\hat{d}_{0:k} \triangleq [\hat{d}_0, \hat{d}_1, \ldots, \hat{d}_k]^T$, $\hat{\theta}_{0:k} \triangleq [\hat{\theta}_0, \hat{\theta}_1, \ldots, \hat{\theta}_k]^T$, $\hat{\alpha}_{0:k} \triangleq [\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_k]^T$ and $\hat{P}_{0:k} \triangleq [\hat{P}_0, \hat{P}_1, \ldots, \hat{P}_k]^T$. Then $F(P_{0:\,k})$ can be given as

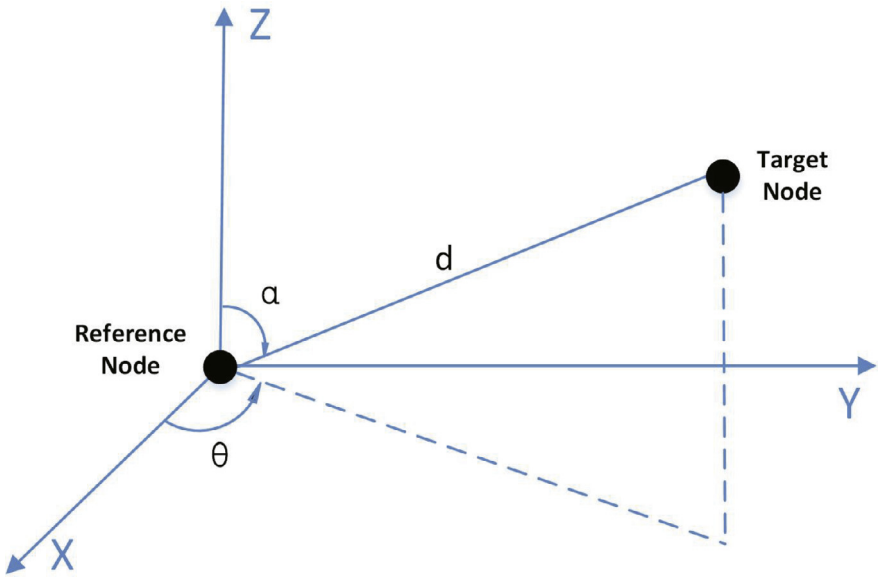


Fig. 3. 3D space parameters between target node and reference node are defined: distance measurement d, azimuth angle θ and pitch angle α.

$$F(P_{0:k}) = \begin{bmatrix} E(-\nabla^{P_{0:k-1}}_{P_{0:k-1}} lnp_k) & E(-\nabla^{P_{0:k}}_{P_{0:k-1}} lnp_k) \\ E(-\nabla^{P_{0:k-1}}_{P_{0:k}} lnp_k) & E(-\nabla^{P_{0:k}}_{P_{0:k}} lnp_k) \end{bmatrix} \triangleq \begin{bmatrix} A_k & B_k \\ B_k^T & C_k \end{bmatrix} \tag{15}$$

and the submatrix F_k can be obtained by block matrix pseudoinverse as

$$F_k = C_k - B_k^T A_k^{-1} B_k \tag{16}$$

Also, the joint p.d.f for *(k+1)*th step p_{k+1} can be written as

$$p_{k+1} = p_k p(\hat{d}_k|P_{k+1}, P_k) p(\hat{\theta}_k|P_{k+1}, P_k) p(\hat{\alpha}_k|P_{k+1}, P_k) p(\hat{r}_k|P_{k+1}) \tag{17}$$

With p_{k+1}, the FIM at *(k+1)*th step $J(P_{0:k+1})$ can be calculated with the same approach as

$$F(P_{0:k+1}) = \left[\begin{array}{cc|c} A_k & B_k & 0 \\ B_k^T & C_k + H_k^{11} & H_k^{12} \\ \hline \mathbf{0} & H_k^{12} & \mathcal{P}_{k+1} + H_k^{22} \end{array}\right] \tag{18}$$

where $\boldsymbol{0}$ is zero matrices with proper dimension. H_k^{11}, H_k^{12}, H_k^{12}, and H_k^{22} carries the posterior information from the difference between $(k+1)$th and kth step, which is referred to as the knowledge of step length, heading and elevation measurements. They can be written as

$$H_k^{11} = E_{\hat{d},\hat{\theta},\hat{\alpha}}[-\nabla^{P_k}_{P_k} lnp(\hat{d}_k|P_{k+1}, P_k) p(\hat{\theta}_k|P_{k+1}, P_k) p(\hat{\alpha}_k|P_{k+1}, P_k)] \tag{19}$$

$$H_k^{12} = E_{\hat{d},\hat{\theta},\hat{\alpha}}[-\nabla^{P_{k+1}}_{P_k} lnp(\hat{d}_k|P_{k+1}, P_k) p(\hat{\theta}_k|P_{k+1}, P_k) p(\hat{\alpha}_k|P_{k+1}, P_k)] = H_k^{21} \tag{20}$$

$$H_k^{22} = E_{\hat{d},\hat{\theta},\hat{\alpha}}[-\nabla^{P_{k+1}}_{P_{k+1}} lnp(\hat{d}_k|P_{k+1}, P_k) p(\hat{\theta}_k|P_{k+1}, P_k) p(\hat{\alpha}_k|P_{k+1}, P_k)] \tag{21}$$

Note that $\mathcal{P}_{k+1}$ denotes to the knowledge of range-based RF localization for joint motion, which is given as

$$\mathcal{P}_{k+1} = E_{\hat{r},P}[-\nabla^{P_{k+1}}_{P_{k+1}}] lnp(\hat{r}_{k+1}|P_{k+1}) \tag{22}$$

The knowledge of RF-based location estimation can be also derived from the inaccuracy of RF-based range estimation as $\mathcal{P} = \mathcal{G}_k^T \Lambda_k^{-1} \mathcal{G}_k$, where $\mathcal{G}_k = \nabla^T_{P_k} \otimes r_k$ and the diagonal matrix $\Lambda_k = \text{diag}\{\lambda_{k,1}, \ldots, \lambda_{k,N}\}$ is introduced to collect the variance of pairwise range estimates with different on-body receivers at the kth step. Note that $r_k = [r_{k,1}, \ldots, r_{k,N}]$ denotes to the vector of actual distance between joints, $\lambda_{k,n}$ denotes to the variance of range estimate between joints. Both of the definitions have been mentioned in previous sections.

With the FIM for entire *(k+1)* steps $J(P_{0:k+1})$ in [(17)], the posterior information submatrix for estimating ukt1 reads

$$F_{k+1} = \mathcal{P}_{k+1} + H_k^{22} - \begin{bmatrix} 0 & H_k^{21} \end{bmatrix} \begin{bmatrix} A_k & B_k \\ B_k^T & C_k + H_k^{11} \end{bmatrix} \begin{bmatrix} 0 \\ H_k^{12} \end{bmatrix} = \mathcal{P}_{k+1} + H_k^{22} - H_k^{21}(F_k + H_k^1 1)(-1) H_k^{12} \tag{23}$$

Since the noises of step length, heading and elevation measurements n_k, u_k and v_k are modeled as zero-mean Gaussian random variables, H_k^{11}, H_k^{12} and H_k^{22} can be calculated as

$$H_k^{11} = -H_k^{12} = H_k^{22} = \mathcal{H}_k \tag{24}$$

where

$$\mathcal{H}_k = \frac{\mathcal{J}_{\hat{d}}(\theta_k, \alpha_k)}{\sigma_k^2} + \frac{\mathcal{J}_{\hat{\theta}}(\theta_k, \alpha_k)}{\epsilon_k^2 sin^2(\alpha) d^2} + \frac{\mathcal{J}_{\hat{\alpha}}(\theta_k, \alpha_k)}{\xi_k^2 d^2} \tag{25}$$

Note that $\mathcal{J}_a(b)$ denotes the numerator of Jacobian matrix, which can be written as

$$\mathcal{J}_{\hat{d}}(\theta_k, \alpha_k) = \begin{bmatrix} \cos^2\theta \sin^2\alpha & \sin\theta\cos\theta\sin^2\alpha & \cos\theta\sin\alpha\cos\alpha \\ \sin\theta\cos\theta\sin^2\alpha & \sin^2\theta\sin^2\alpha & \sin\theta\sin\alpha\cos\alpha \\ \cos\theta\sin\alpha\cos\alpha & \sin\theta\sin\alpha\cos\alpha & \cos^2\alpha \end{bmatrix} \tag{26}$$

$$\mathcal{J}_{\hat{\theta}}(\theta_k, \alpha_k) = \begin{bmatrix} \sin^2\theta & -\sin\theta\cos\theta & 0 \\ -\sin\theta\cos\theta & \cos^2\theta & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{27}$$

$$\mathcal{J}_{\hat{\theta}}(\theta_k, \alpha_k) = \begin{bmatrix} \cos^2\theta\cos^2\alpha & \sin\theta\cos\theta\cos^2\alpha & -\cos\theta\sin\alpha\cos\alpha \\ \sin\theta\cos\theta\cos^2\alpha & \sin^2\theta\cos^2\alpha & -\sin\theta\sin\alpha\cos\alpha \\ -\cos\theta\sin\alpha\cos\alpha & -\sin\theta\sin\alpha\cos\alpha & \sin^2\alpha \end{bmatrix} \tag{28}$$

Therefore, the posterior information submatrix F_{k+1} for estimating the joint's location P_{k+1} can be recursively calculated by

$$F_{k+1} = \mathcal{P}_{k+1} + \mathcal{H}_k - \mathcal{H}_k(F_k + \mathcal{H}_k)^{-1}\mathcal{H}_k \tag{29}$$

Applying Sherman Morrison Woodbury formula with identity matrices of proper dimension, it can be simplified as

$$F_{k+1} = \mathcal{P}_{k+1} + (\mathcal{H}_k^{-1} - F_k^{-1})^{-1} \tag{30}$$

where $\mathcal{H}_k$ shows the effect of image processing based step, heading and elevation measurements and $\mathcal{P}_{k+1}$ illustrates the effect of RF based location estimation.

3.3. Performance

In this section, the hybrid localization approach has been implemented and the PCRLB has been calculated to evaluate the accuracy of the proposed approach. Even though there are elegant expressions to recursively calculate the FIM, the expressions in Eq. (29) usually do not have analytical close-form solution. In order to deal with that, we employ the monte carlo approach to convert those continuous integrals into discrete summations, and finally work out the PCRLB.

The Root-Mean-Square of PCRLB is given by $\frac{1}{\mathcal{I}}\sum_{i=1}^{\mathcal{I}}\mathbf{\Theta}_k^i$, where $\mathbf{\Theta}_k^i$ is the PCRLB on the Root-Mean-Square Error (RMSE) of joint at step k in the ith *monte carlo* trial. i denotes the index for monte carlo trials and $\mathcal{I}$ is the total monte carlo trial number ($\mathcal{I} = 2000$ is used in this paper). The sampling rate is set as 50 Hz, and horizonal and vertical velocity is both 1 m/s. Note that for each monte carlo trial, we randomly select the initial location for PCRLB calculation in order to get a fair average of entire human body moving process.

The experimental scenario is given as follows. Inaccuracy on step length measurement are considered to be proportion of the actually step length d as $\sigma_k = \eta d_k$, heading and elevation measurements are considered to have identical variance as $\epsilon_k = \xi_k = \omega$ are calculated with η varying from 5% to 10% and ω varying from 10° to 30° to illustrate the minimum achievable RMSE $\mathbf{\Theta}_k$ in hybrid joint localization. PCRLB of the method only using inertial sensors or RF sensors have been also calculated to provide a comparison against PCRLB.

Results on PCRLB with various η and ω have been plotted in Fig. 4, in which the PCRLB with the absence of inertial or distance sensing are also provided for comparison. It can be seen that (1) The knowledge of step length, heading and elevation measurements significantly increase the accuracy of joint capturing. The PCRLB for hybrid localization drops below 8 cm. (2) The hybrid localization accuracy is directly proportional with the accuracy of step length, heading and elevation measurements, that is, with the increment of η and ω, obvious decrement of RMSE can be seen from the PCRLB plot. (3) PCRLB for hybrid localization stabilizes after certain steps regardless of the beginning point and following trajectory. With better posterior knowledge of step length and headings, the PCRLB stabilize slower. However, with poor step length, heading and elevation measurements, the hybrid localization approach reaches the maximum achievable performance very fast. (4) Distinguished from hybrid localization methods, if only inertial sensors are used in the monitoring system, the error may tend to be diverging. In theory, this confirms that inertial capturing system faces the problem of accumulated errors. Besides, it shows that our hybrid method can avoid the divergence of accumulated errors.

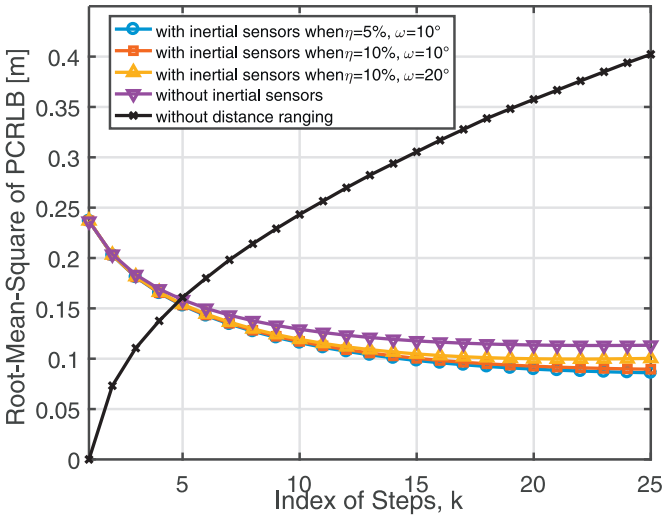


Fig. 4. PCRLB of the proposed D–H based human motion capturing approach as a function of step index.

Different combination of sensors and anchors are also considered, which potentially represent the Anchor sensor arrangement in practical applications. As described in Section 2, the trunk remains relatively steady in comparison to the end joints or limbs of the human body. In consideration of end-effector problem, we set the joints of neck, chest, left and right hip as reference nodes for error bounds estimation. A scenario squared by $2\text{m} * 2\text{m}$ is chosen as test field. The experiment human body was placed in the center of this field and the above reference nodes are respectively located at joints in the trunk. For comparison, four sets of reference nodes combination were chosen serving as a contrast. The selected candidate topologies is shown in Fig. 5.

PCRLB of above mentioned topologies are displayed in Fig. 6, from which we can see that (1) Comparing between different candidate topologies, the improvement on localization accuracy remains identical for different topologies. (2) The hybrid localization with different topologies shares the same number of steps to stabilize. (3) Topology 3 suffers the largest RMSE, may be due to the dense distribution of candidate Anchors. Topology 4 is preferred for its minimum RMSE, which could be less than 8 cm. The hybrid location estimation with RF ranging shares identical trend results with that of solely RF ranging.

From above results, the effectiveness of our proposed geometrical model is verified. Due to the end-effect of human motion, our model takes D–H parameters into consideration which contains both distance and IMU information. The relative position of different human body parts is also considered. These superiorities contribute to the lower PCRLB and higher capture accuracy.

4. Model evaluation

The theoretical effectiveness of our proposed motion model is verified in above section. The introduction of D–H parameters utilizes both distance and IMU information to lower the expected PCRLB in theory. Furthermore, we take advantage of this method in practical application as a validation.

Human lower limb motion analysis is a typical application of wearable sensing platform. Traditional systems all tend to monitor foot phase when walking, which just reflects the rhythm of footsteps. However, ignoring the relative position between other body

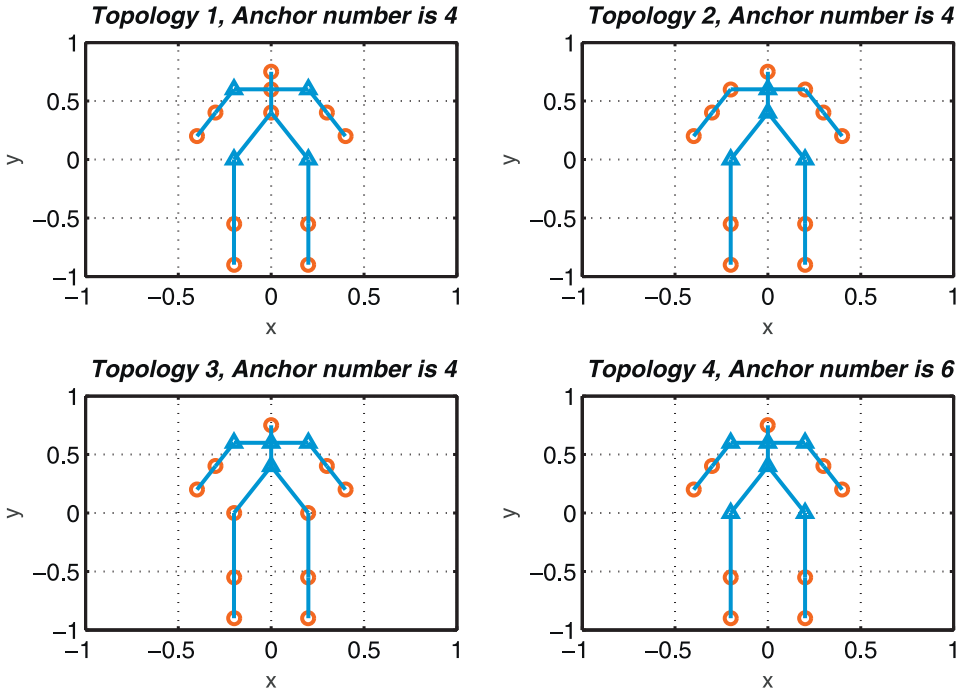


Fig. 5. Topologies configuration. ○ represents free joints and △ represents selected anchors.

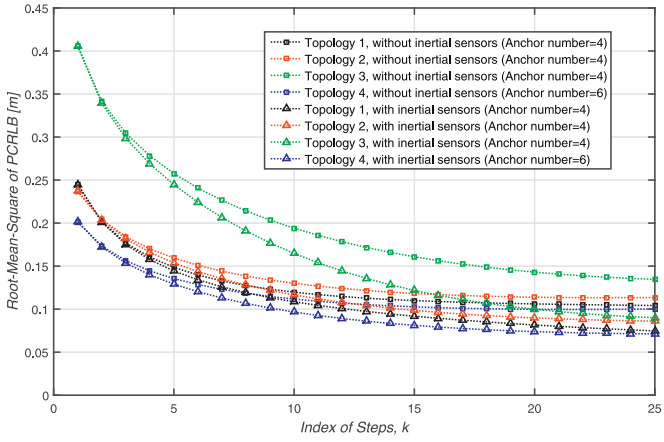


Fig. 6. Root-mean-square of PCRLB as a function of index of steps. Different receiver topologies are considered.

parts, especially leg swings, leads to the lack of some important features and effective constraint conditions. With the help of proposed geometrical model, using D-H convention, we can propose limb transformations from the ankle joint to the hip joint. Motion characteristics of various body parts can easily be obtained, and we may take this analysis as an example to verify the effectiveness.

4.1. Platform overview

In order to describe the details of human motions, motion transition matrix M_0^k needs to be constructed to represent the spatial and temporal features. For this reason, a geometrical kinematics based wearable sensing platform is designed. The platform is aiming to collect spatial and temporal information (namely D–H parameters) during the human motion process. Sensing nodes are designed and intended to be put on joints to capture the movement conditions of joints as well as limbs. The information covers accelerated velocity, angular velocity and distances between joints and body parts. Among these all, distance information is especially special when compared with other platforms. Experiments are conducted on the data set sampled by the above platform at 10 Hz. We use the presented platform for data collection and perform all processing work online in Matlab with PC (Intel Core i7-6700M CPU, 16GB RAM). The communication between wearable nodes and the PC is via Bluetooth.

Our integrated wearable sensor system is composed of two parts: one control unit and several data acquisition units. The control unit works as a gateway to control the whole operation process via Bluetooth communication. Data acquisition unit is mainly responsible for data sensing. Each data acquisition unit has a 6-axises sensor (MPU6050, which integrates a triaxial accelerometer and a triaxial gyroscope), a barometer sensor (MS5611) and a UWB TOA ranging module (DWM1000). The MEMS sensors are connected to a micro-controller (STM32F103) for the sake of sampling efficiency in a rate of 10 Hz. Data are transferred to the control unit in real-time and also written the realtime data stream into its SD card as backups for other offline analysis and applications. The whole system architecture is demonstrated in Fig. 7.

4.2. Experiment setup

In common sense, a whole lower limb movement cycle could be divided into two phases: stance phase and swing phase. Furthermore, several fine-grained division are studied in [6], each part of which owns its related features and parameters. According to related references, zero velocity update (ZUPT) [6,7] is the most widely adopted method to cope with sensor drift errors when the foot is at rest on the ground in stance phase. However, it should be noted that ZUPT method is valid only when zero velocity assumption is satisfied. Therefore, ZUPT is invalid for knee position estimation because there is no static moment for the knee during walking.

Likewise, distortion could happen toward the motion tracking of other parts of human body other than feet. The articulated chain structure of proposed human model (especially for lower limbs) makes it applicable to set up geometrical constraints [30], which

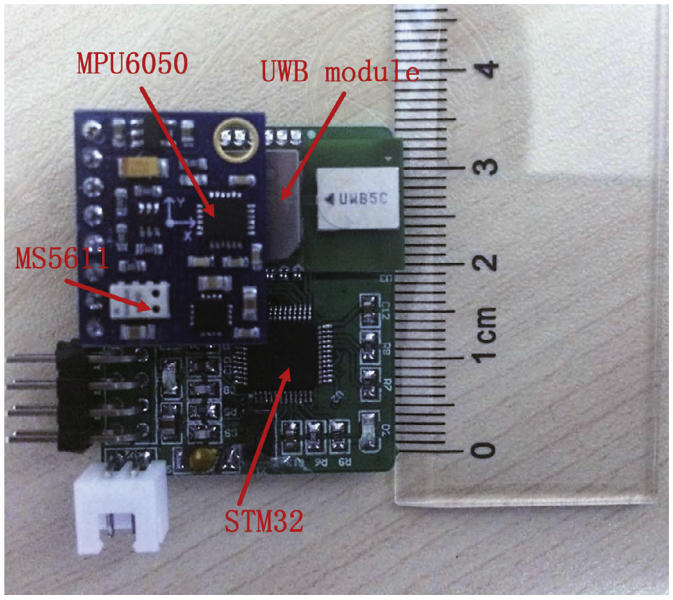


Fig. 7. Experimental Platform Settings.

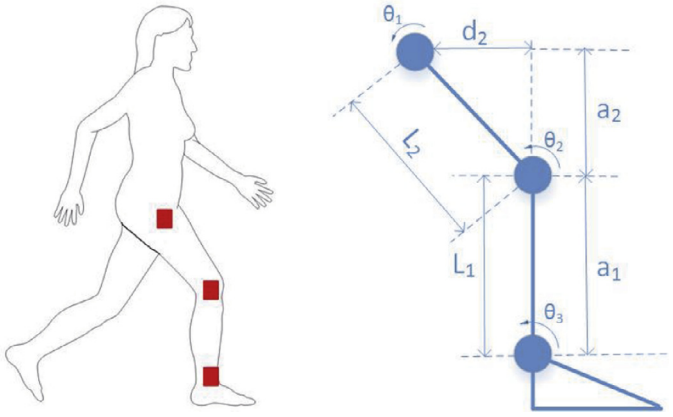


Fig. 8. Placements deployed with sensing devices and parameters of D-H needed to be measured.

is based on the concept of forward kinematics from the robot domain. For the above reasons, our proposed geometrical human motion model is applied. With the introducing of D–H convention, joints position are more easily to be obtained with considering linked body parts. Fluctuant errors can be somehow avoided with cascaded constraints taken into consideration. Three sensor nodes are separately placed on different parts of lower limbs to form constrained states with the complementary geometric relation, which offers us a way to correct lower limb motion errors [31].

For above considerations, three monitoring devices are deployed on the joints of right leg, namely right ankle, right knee and right hip, shown in Fig. 8. Since human lower limbs can be treated as a model of inverted pendulum, the geopotential energy can reach the maximum at the minimum of kinetic energy (the vertical moment), which is the start time of the implementation of D–H convention. Knee position estimation [32] is evaluated by combining the accurate position information of the foot and the D–H convention.

4.3. Capturing accuracy

Suppose that all the sensors were kept still before the sampling starts and the displacement of each sensor is relevant to its original position. The inverted pendulum structure of legs significantly makes hip joint more suitable to be root node when applied D–H convention. The variations of the knee at time i, denoted as $P_i^k = \begin{bmatrix} P_{i,x}^k & P_{i,y}^k & P_{i,z}^k \end{bmatrix}^T$, derived from the displacement vector of the hip joint $P_i^h = \begin{bmatrix} P_{i,x}^h & P_{i,y}^h & P_{i,z}^h \end{bmatrix}^T$ and the basis of motion vector generates P_i^h is the displacement vector of the foot, $P_i^f = \begin{bmatrix} P_{i,x}^f & P_{i,y}^f & P_{i,z}^f \end{bmatrix}^T$. These transitive relation then forms two constraint states.

$$P_i^k = P_i^h + R_{k-1}^k P_{i-1}^h \tag{31}$$

$$P_i^f = P_i^k + R_{k-1}^k P_{i-1}^k \tag{32}$$

In this way, we can substantially confine the broadening of drift errors with proposed geometrical model, considering physical relationship when human body conducts movements. To verify the effectiveness of this proposed method in human lower limb analysis, we experimented with human walking scenarios with designed modules. One of the subject's foot walking trajectory is shown in Fig. 9. A sequence of movements lasted about 45 seconds, including climbing spiralstairs and part of walking on level ground.

For comparison, both tracking trajectory, with or without, using the proposed method are drawn in Fig. 9. It is clearly seen that, with applying the geometrical constrained model, the experiment results are far closer to the ground truth, while on the other hand, the results without constraints are drifting away as time accumulates. The results remain similar at the very beginning; however, the gap became larger while time goes on. Detailed accumulated

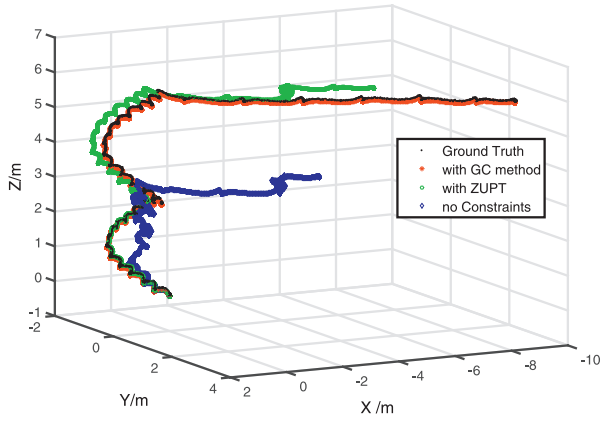


Fig. 9. Walking trajectory when climbing a spiralstairs with or without applying GC method.

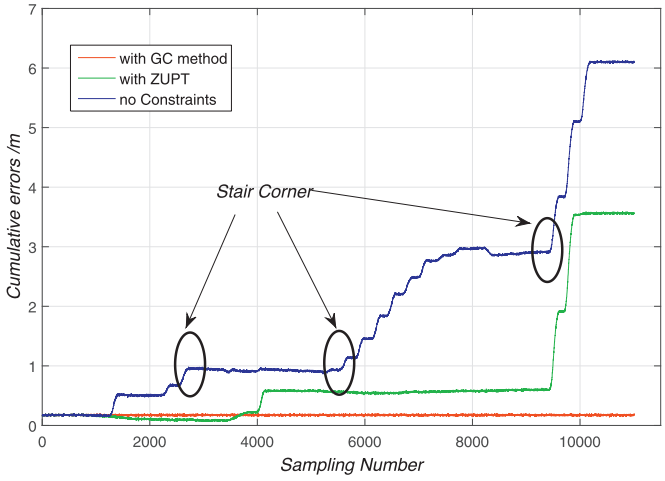


Fig. 10. Comparison of navigation errors as time cumulates.

Table 1
Comparison of different capturing algorithms.

Method	Computation cost [mW]	Communication cost [mW]	Total consumption [mW]
No constraints	99.7	0	99.7
ZUPT	138.7	0	138.7
GC method	133.06	0.6	133.66

errors are shown in Fig. 10. Methods with or without using ZUPT both face the problem of drift error, especially when the experimenter turns around at stair corners. Time accumulated errors are more likely to be caused when moving direction changes. With ZUPT method applied into the algorithm, localization error to some extent can be fixed, but it still exists and it is also critical at turns. On the contrary, our proposed human motion model shows a good stability and no clearly drift errors are observed in our experiment. This is because with the use of proposed model, body connections are considered and temporal transitions are estimated. Especially at turns, body constrains can to some extent fix time accumulated errors. The results suggest that our proposed GC method performs well while applied in human lower limb analysis. This also confirms the conclusion obtained in PCRLB simulation. Methods with only inertial sensors face the problem of drift errors, but our geometrical constrained model can fix time accumulated errors.

4.4. Energy consumption

In this paper, we intend to propose a method that can obtain an as high accuracy as possible, with acceptable power consumption. Compared with traditional methods, our proposed model can achieve a rather high motion capturing accuracy, which can be seen from Fig. 10. Thus, we further focus on the assessment of energy consumption. To calculate consumption of our system, we consider two sources of energy consumption, namely computation and communication costs [34], denoted as P_{comp} and P_{comm} respectively. The power consumption of the system due to data acquisition and analysis is then given by:

$$Z = P_{comp} + P_{comm} \tag{33}$$

1) ***Computation cost:*** It mainly comes from several aspects, including sensor power consumption, and the computation cost for combining results of distributed sensors.
   - Sensor power consumption: MPU6050 [35], which integrated acceleration and gyroscope, is chosen to measure the information needed by our proposed model, shown in Eq. (3). The nominal values of power used by MPU6050 is no more than 14.1 mW [35]. Given that data acquisition from a sensor (i.e., accelerometer or gyroscope) is constant at a fixed sampling rate of 10 Hz, we allocate a fixed cost to each sensor. As a result, the sensor power consumption of each motion node is equal to 14.1 mW.
   - Algorithm computation cost: MCU [36] computation cost is the most important part in the analysis of energy consumption. In our method, we use an event-driven methodology to reduce communication traffic. At the initial state, each device goes to sleep mode to minimize energy and is waiting for timer interrupt events. Once the interrupt occurs, MCU wakes up from the sleep mode, starts to capture sensor signal and transmits it to central unit. The central unit runs our proposed modeling based method. In order to balance energy consumption, we switch central node periodically.
2) ***Communication cost:*** It mainly comes from the communication overhead when the wearable sensors exchange data with each other. Each captured data is immediately processed and transmitted in batch periodically, and the sensor node turns into sleep mode back . The nominal values of UWB transmitter is 231 mW in Tx mode, 99 mW in Rx mode, denoted as P_{Tx} and P_{Rx} respectively. Transmitted package is in fixed length (raw data in each cycle). Generally, the maximum transmission rate of dw1000 can reach 6.8Mbps [37], transmitting time (t_{Tx}) is in nanoseconds, and receiving time (t_{Rx}) is about 2 ms. Thus, the working consumption could be calculated by

$$P_{working} = P_{Tx} * t_{Tx}/(t_{Tx} + t_{Rx}) + P_{Rx} * t_{Rx}/(t_{Tx} + t_{Rx}) \tag{34}$$

Its sleep current can be as low as 50 nA, when 3.3 V supplies [37]. The duty cycle (T_{duty}) can be very low ($<3/1000$). Thus, power cost by ranging and data transmitting is almost constant, denoted as

$$P_{comm} = P_{working} * T_{duty} + P_{sleep} * (1 - T_{duty}) \tag{35}$$

which is around 0.6 mW in our system.

For comparison, we measured the power consumption of the three algorithms, shown in Table 1. In the aspect of computation cost, ZUPT and our proposed GC method are slightly higher than no constraints condition, as they have higher computational complexity. On the other hand, for communication cost, only our proposed method needs to transmit data between nodes. However, due to the low duty cycle, its average power consumption is relatively low. Overall, GC method have slightly less power consumption than ZUPT. This is because with the use of proposed model, the calculation is almost all linear transformations. It makes use of communication between nodes, whose power consumption is low, to integrate multiple data sources. As a result, it can reduce the complexity of the algorithm.

From above analysis, GC method has significantly higher accuracy, as well as acceptable power consumption. It is more suitable for wearable motion capturing applications.

5. Conclusion

In this paper, we put forward a geometrical kinematic characteristics based human motion model. We view the whole human body as a connected articulated entirety and modeled human motion with dynamic kinematics analysis using Denavit–Hartenberg Convention. The complementary geometric relation between joints and limbs is considered in human motion sensing process, which to some extent limits the sensor drift error. PCRLB of human motion process are derived based on proposed geometrical model. Simulation results show superior performance when proposed model is utilized with considering both distance and IMU information. As field experiment, the model is applied in human lower limb motion sensing applications. Significant reduction in the localization error and acceptable energy consumption are shown in the testing results, which is much more efficient than traditional method ZUPT.

Acknowledgment

This work is supported by National Natural Science Foundation of China (NSFC) project No. 61671056, No. 61302065, No. 61304257, No.61402033, Beijing Natural Science Foundation project No.4152036 and Tianjin Special Program for Science and Technology No. 16ZXCXSF00150.

References

[1] P.N. Pathirana, S. Li, H.M. Trinh, et al., Robust real-time bio-kinematic movement tracking using multiple kinects©for tele-rehabilitation, IEEE Trans. Ind. Electron. (2016). 1–1

[2] G. Fortino, R. Giannantonio, R. Gravina, et al., Enabling effective programming and flexible management of efficient body sensor network applications, IEEE Trans. Hum. Mach. Syst. 43 (1) (2012) 115–133.

[3] R. Gravina, P. Alinia, H. Ghasemzadeh, G. Fortino, Multi-sensor fusion in body sensor networks: state-of-the-art and research challenges, Inf. Fusion 35 (2017) 68–80.

[4] S. Iyengar, F.T. Bonda, R. Gravina, et al., A framework for creating healthcare monitoring applications using wireless body sensor networks, in: Proceedings of the ICST 3rd International Conference on Body Area Networks. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

[5] Y. Chen, B. Lu, Y. Chen, X. Feng, Breathable and stretchable temperature sensors inspired by skin, Sci. Rep. 5 (2015) 11505. EP C, 06 2015. [Online]. Available: doi: 10.1038/srep11505

[6] O. Bebek, et al., Personal navigation via high-resolution gait-corrected inertial measurement units, IEEE Trans. Instrum. Meas. 59 (11) (2010) 3018C3027.

[7] S.J.M. Bamberg, A.Y. Benbasat, D.M. Scarborough, D.E. Krebs, J.A. Paradiso, Gait analysis using a shoe-integrated wireless sensor system, IEEE Trans. Inf. Technol. Biomed. 12 (4) (2008) 413C423.

[8] H. Ghasemzadeh, R. Jafari, Physical movement monitoring using body sensor networks: a phonological approach to construct spatial decision trees, IEEE Trans. Ind. Inform. 7 (1) (2011) 66–77.

[9] S. Zihajehzadeh, P.K. Yoon, B.S. Kang, et al., UWB-aided inertial motion capture for lower body 3-d dynamic activity and trajectory tracking, IEEE Trans. Instrum. Meas. 64 (12) (2015) 3577–3587.

[10] H. Luinge, P. Veltink, Measuring orientation of human body segments using miniature gyroscopes and accelerometers, Med. Biol. Eng. Comput. 43 (2) (2005) 273C282.

[11] G. Ligorio, E. Bergamini, I. Pasciuto, et al., Assessing the performance of sensor fusion methods: application to magnetic-inertial-based human body tracking, Sensors 16 (2) (2016).

[12] S. Liu, R.X. Gao, D. John, et al., Multisensor data fusion for physical activity assessment, IEEE Trans. Biomed. Eng. 59 (3) (2012) 687–696.

[13] Y. Tao, H. Hu, A novel sensing and data fusion system for 3-d arm motion tracking in telerehabilitation, IEEE Trans. Instrum. Meas. 57 (5) (2008) 1029C1040.

[14] G. Wu, S. Siegler, P. Allard, et al., ISB Recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part i: ankle, hip, and spine, Int. Soc. Biomech. J. Biomech. 35 (4) (2002) 543–548.

[15] G. Wu, V.D.H. Fc, H.E. Veeger, et al., ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part II: shoulder elbow, wrist and hand, J. Biomech. 38 (5) (2005) 981–992.

[16] D. Roetenberg, H. Luinge, P. Slycke, MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors, Xsens Motion Technologies Bv, 2009.

[17] E. Foxlin, M. Harrington, Weartrack: a self-referenced head and hand tracker for wearable computers and portable VR, in: Proceedings of the 4th International Symposium Wearable Computation, 2002, p. 155C162.

[18] D. Demirdjian, Enforcing constraints for human body tracking, in 2003, in: Conference on Computer Vision and Pattern Recognition Workshop Vol. 9, Madison, Wisconsin, USA, 2003, p. 102C109.

[19] E.A. Hunter, P.H. Kelly, R.C. Jain, Estimation of articulated motion using kinematically constrained mixture densities, in: Proceedings of the Nonrigid and Articulated Motion Workshop, 1997, IEEE, 1997, pp. 10–17.

[20] S. Knoop, S. Vacek, R. Dillmann, Modeling joint constraints for an articulated 3d human body model with artificial correspondences in ICP, in: IEEE-RAS International Conference on Humanoid Robots, IEEE, 2005, pp. 74–79.

[21] M. Mihelj, Inverse kinematics of human arm based on multisensor data integration, J. Intell. Robot. Syst. 47 (2) (2006) 139–153.

[22] D. Vlasic, R. Adelsberger, G. Vannucci, et al., Practical motion capture in everyday surroundings, ACM Trans. Graph. 26 (3) (2007) 35.

[23] Y. Feng, C. Fritsche, F. Gustafsson, et al., TOA-Based robust wireless geolocation and cramr-rao lower bound analysis in harsh LOS/NLOS environments, IEEE Trans. Signal Process. 61 (9) (2013) 2243–2255.

[24] Y. Wu, X. He, B. Kang, et al., Long-range motion trajectories extraction of articulated human using mesh evolution, IEEE Signal Process. Lett. 23 (4) (2015) 507–511.

[25] J. Kim, M. Kim, Articulated human motion tracking using a SOG-PF algorithm, in: International Conference on Information and Communication Technology Convergence, IEEE, 2015.

[26] V.J. Santos, F.J. Valero-Cuevas, Anatomical variability naturally leads to multimodal of denavit-hartenberg parameters for the human thumb, IEEE Trans. Biomed. Eng. 53 (2) (2006) 155–163.

[27] G. Welch, G. Bishop, An Introduction to the Kalman Filter, University North carolina at chapel Hill, 1995.

[28] J.C. Spall, Monte carlo computation of the fisher information matrix in nonstandard settings, J. Comput. Graph. Stat. 5 (4) (2012) 889–909.

[29] J. Barraquand, J.C. Latombe, Robot motion planning: a distributed representation approach, Int. J. Rob. Res. 10 (6) (1991) 628–649.

[30] Z.-Q. Zhang, W.-C. Wong, J.-K. Wu, Ubiquitous human upper-limb motion estimation using wearable sensors, IEEE Trans. Inf. Technol. Biomed. 15 (4) (2011) 513C521.

[31] T. Plötz, N.Y. Hammerla, P. Olivier, Feature Learning for Activity Recognition in Ubiquitous Computing, in: IJCAI 2011, Proceedings of the, International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, 2011, pp. 1729–1734.

[32] X. Meng, Z.-Q. Zhang, J.-K. Wu, W.-C. Wong, H. Yu, Selfcontained pedestrian tracking during normal walking using an inertial/magnetic sensor module, IEEE Trans. Biomed. Eng. 61 (3) (2014) 892C899.

[33] P. Tichavsky, C.H. Muravchik, A. Nehorai, Posterior cram!'äer-rao bounds for discrete-time nonlinear filtering, IEEE Trans. Signal Process. 46 (5) (1998) 1386C1396.

[34] H. Ghasemzadeh, P. Panuccio, S. Trovato, et al., Power-aware activity monitoring using distributed wearable sensors, IEEE Trans. Hum. Mach. Syst. 44 (4) (2014) 537–544.

[35] InvenSense website. (2013). [Online]. Available: https://www.invensense.com/.

[36] STMicroelectronics website. (2013). [Online]. Available: http://www.st.com/.

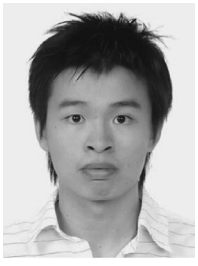
[37] DecaWave website. (2013). [Online]. Available: https://www.decawave.com/.

Cheng Xu received the B.E. and M.S. degree from the University of Science and Technology Beijing (USTB), China in 2012 and 2015, respectively. He is currently working towards the PhD degree in the Data and Cyber-Physical System Lab (DCPS) at University of Science and Technology Beijing. His research interests include patter recognition and internet of things. He is a student member of the IEEE and CCF.

Jie He received B.E. and Ph.D degree in computer science from University of Science and Technology Beijing (USTB), China in 2005 and 2012, respectively. Since July 2015, he has been an associate professor with the School of Computer and Communication Engineering, USTB since 2015. From April 2011 to April 2012, he was a visiting Ph.D student in Center for Wireless Information Network Studies, Worcester Polytechnic Institute. His research interests include wireless indoor positioning, human gesture recognition and motion capture.

Xiaotong Zhang received the M.S., and Ph.D.degrees from University of Science and Technology Beijing, in 1997, and 2000, respectively. He was a Professor in the Department of Computer Science and Technology, University of Science and Technology Beijing. His research includes work in quality of wireless channels and networks, wireless sensor networks, networks management, cross-layer design and resource allocation of broadband and wireless network, signal processing of communication and computer architecture.

Cui Yao received the B.E. degrees from the University of Science and Technology Beijing, Beijing, China, in 2014 and he is currently working toward the Master degree in the Data and Cyber-Physical System Lab (DCPS) at University of Science and Technology Beijing. His research interests include patter recognition and internet of things.



Po-Hsuan Tseng received the B.S. and Ph.D. degrees from the National Chiao Tung University in 2005 and 2011, respectively. He has been an associate professor with National Taipei University of Technology (NTUT), Taipei, Taiwan since 2017. He was a visiting researcher with the University of California at Davis in 2010. His research interests are in the areas of signal processing for networking and communications. He is an honorary member of the Phi Tau Phi Scholastic Honor Society of R.O.C. He is a member of the IEEE.