

Article

Towards Human Motion Tracking: Multi-Sensory IMU/TOA Fusion Method and Fundamental Limits

Cheng Xu ^{1,2} , Jie He ^{1,2,*}, Xiaotong Zhang ^{1,2,*} , Xinghang Zhou ^{1,2} and Shihong Duan ^{1,2}

¹ School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China; xucheng19880202@foxmail.com (C.X.); zhouxinghang@xs.ustb.edu.cn (X.Z.); duansh@ustb.edu.cn (S.D.)

² Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

* Correspondence: hejie@ustb.edu.cn (J.H.); zxt@ies.ustb.edu.cn (X.Z.)

Received: 18 January 2019; Accepted: 27 January 2019; Published: 29 January 2019



Abstract: Human motion tracking could be viewed as a multi-target tracking problem towards numerous body joints. Inertial-measurement-unit-based human motion tracking technique stands out and has been widely used in body area network applications. However, it has been facing the tough problem of accumulative errors and drift. In this paper, we propose a multi-sensor hybrid method to solve this problem. Firstly, an inertial-measurement-unit and time-of-arrival fusion-based method is proposed to compensate the drift and accumulative errors caused by inertial sensors. Secondly, Cramér–Rao lower bound is derived in detail with consideration of both spatial and temporal related factors. Simulation results show that the proposed method in this paper has both spatial and temporal advantages, compared with traditional sole inertial or time-of-arrival-based tracking methods. Furthermore, proposed method is verified in 3D practical application scenarios. Compared with state-of-the-art algorithms, proposed fusion method shows better consistency and higher tracking accuracy, especially when moving direction changes. The proposed fusion method and comprehensive fundamental limits analysis conducted in this paper can provide a theoretical basis for further system design and algorithm analysis. Without the requirements of external anchors, the proposed method has good stability and high tracking accuracy, thus it is more suitable for wearable motion tracking applications.

Keywords: Cramér–Rao lower bound (CRLB); human motion; Inertial Measurement Unit(IMU); Time of Arrival (TOA); wearable sensors

1. Introduction

Human Motion Tracking (HMT) [1,2] has been a hot topic in the area of body area network (BAN) [3,4] during the last decade. It aims to obtain human body movement information via quantitative methods to capture and analyze human motion. The information obtained by HMT are mainly relative positions in 2D/3D space of various joints and limbs. With the booming of BAN [5–7], human-centric applications have been raising in both academic and industrial areas, such as interactive games, computer cinematography, animation, etc. [1]. Many entrepreneurial firms have been formed in these areas [8,9]. Besides, HMT also plays a big role in medical field [5], and military and security applications [6].

HMT systems could be classified into two categories: computer vision (CV)-based [10] and wearable sensor (WS)-based [6]. CV-based technology takes advantages of deployed web-depth cameras to monitor human activities, and has been applied in practical applications, such as film shooting [11], security monitoring [6] and so on. Many technology companies have also been working on the research and development of professional human motion tracking systems, such

as Vicon [12] and Optotrak [13]. They perform with high accuracy when operated in controlled scenario, however they rely heavily on environment factors, such as shooting angle and lights [10]. These inevitable defects make it only suitable for smaller and controllable situations. WS-based HMT systems, by contrast, do not need to deploy devices in scenarios and are less sensitive to the environment. Thus, they are more suitable for large-scale and dynamic applications [14].

In the present stage, WS-based HMT systems are mainly composed of inertial measurement units (IMUs), such as accelerometers and gyroscopes. The basic principle is to measure the triaxial acceleration and angular velocity of the motion by these sensors, and obtain the trajectory of monitoring points through integral operations [15–18]. However, inertial sensors may inevitably throw off errors that accumulate over time [19–21]. The accumulative and drift error is the biggest challenge faced by WS-based HMT systems.

For better solving the drift error problem and improving HMT accuracy, the most common methods are as follows:

1. The hardware aspect uses inertial sensors with high precision, such as XSens [15] and Invensense [16] tracking units.
2. The algorithm aspect enhances the system by using multi-sensor data fusion means [1,19,22,23].

However, the above-mentioned methods cannot fundamentally solve the drift problem. High precision hardware used by wearable tracking systems is usually very expensive, and the cost of a single suit could be hundreds of thousands or even millions of dollars [15–18]. In the aspect of algorithms, filtering methods, such as Kalman and particle filter [24], can somehow slow down the accumulation process, but cannot eliminate it completely. Thus, the inevitable drift problem of inertial sensors has become a crucial constraint for wearable HMT systems, which limits its use in long term or large space applications.

In essence, human motion tracking can be viewed as a local multi-target real-time high-precision three-dimensional positioning problem [23]. Ultra-Wideband (UWB)-based time-of-arrival (TOA) ranging is the most commonly used high-precision localization technology, its measurement accuracy can reach the centimeter level, and it does not have the drawback of accumulative errors, compared with inertial-sensor-based HMT systems [23,25]. The size of TOA chip is small enough to be integrated into wearable devices. Thereinto, IMU/TOA fusion is an effective way to overcome the accumulative errors of drifting problem faced by solo IMU method [26].

Some achievements on IMU/TOA fusion have been reported in many studies (e.g., [20,21,23,25]) However, state-of-the-art studies, (e.g., [23,27]), mainly face the following two drawbacks:

1. Requirements of fixed external anchors: They need to be deployed in certain scenarios; for example, Zihajehzadeh et al. [23] introduced a magnetometer-free algorithm for human lower-body motion tracking by fusing inertial sensors with an UWB localization system. However, it is hard to realize wearable systems, and is not suitable for BAN applications.
2. State-of-the-art studies seldom concentrate on the fundamental limits of IMU/TOA fusion methods, and the error correction effects are not satisfying. For example, Nilsson et al. [27] proposed a cooperative localization method by fusion of dual foot-mounted inertial sensors and inter-agent UWB ranging, but the experiment results show that, compared with the performance lower bounds [22], there is still a lot of room for improvement.

Based on these considerations, we propose an external-anchor-free IMU/TOA fusion method, and analyze its fundamental limits to lay a theoretical foundation for realizing low cost and high precision motion tracking system that is suitable for medium and long term use in large space.

The rest of the paper is organized as follows: Section 2 briefly introduces the related works about human motion tracking techniques and Cramér–Rao lower bound (CRLB). Section 3 describes the problem definition, IMU/TOA fusion-based model and its error sources. Section 4 introduces our external-anchor-free IMU/TOA fusion method, and analyzes its fundamental limits. Sections 5 and 6,

respectively, verify the spatial and temporal performance of proposed fusion-method. Section 7 presents a practical use case to verify the feasibility of proposed method when compared with state-of-the-art methods. Finally, Section 8 presents the conclusion.

2. Related Work

The development of HMT has been rising with the booming of Internet of Things (IoTs) and the rapid progress of human-centric applications [28–31] in the last decades. Particularly, HMT has become an essential task within the fields, such as clinical, military and security applications. In this section, we present a brief literature review on the following aspects: human motion tracking systems and applications, multi-sensory fusion methods and CRLB.

2.1. HMT Systems and Applications

HMT has shown a tremendous potential in the areas of industrial applications. Benefiting from the development of various systems, HMT has permeated into every aspect of social life, including clinical areas, emergency and rescue areas, security and entertainment, to name a few [5–9]. The most widely used HMT systems could be mainly classified into two categories: CV-based [10] and WS-based [6]. Next, we give a brief introduction and discussion about their merits and demerits.

Computer-vision-based systems, such as Vicon [12] or Optotrak [13], have a high accuracy when operating in controlled environments, e.g., several fixed cameras calibrated and correlated in a specific place and capturing configuration. Famous TV shows and movies, such as *The Walking Dead*, *Game of Thrones* and *Guardians of the Galaxy*, are all powered by human motion animators [15]. CV-based systems can provide a large amount of redundant data. Ambulatory systems, such as those using a Kinect [32] to capture human motion, are set in relatively uncontrolled environments and have a restricted field of view. These systems have a restricted margin of maneuverability and are more suitable for indoor use.

In contrast, a very promising frontier for reliable human motion tracking is WS-based [6] system, especially using IMUs. Iyengar et al. [33] defined a framework that managed common tasks for healthcare monitoring applications to aid development of BAN. Anwary et al. [34] utilized a pressure sensor array and IMU for gait analysis. Ghasemzadeh et al. [35] introduced a novel classification model that identified physical movements from body-worn inertial sensors while considering the collaborative nature and limited resources of the system. Inertial sensors are integrated with UWB localization system in [23] for simultaneous 3-D trajectory tracking and lower body motion capture (MoCap) under various dynamic activities such as walking and jumping.

WS-based HMT systems do not need to deploy devices in scenarios and are less sensitive to the environment. Thus, they are more suitable for large-scale and dynamic applications [2,6]. Currently, WS-based HMT systems [6] are mainly composed of inertial sensors, such as accelerometers and gyroscopes. However, inertial sensors may inevitably throw off errors that accumulate over time [8,15,25]. The accumulative and drift error is the biggest challenge faced by WS-based HMT systems.

2.2. Sensor Fusion and Filtering

Despite the advantages mentioned above, there still exist challenges when wearable sensors are applied to human motion tracking. As human motion tracking could be viewed as a multiple-target 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 [8,15,25].

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. Yun et al. [36] realized a method to measure the orientation of human body segments using Kalman filter that takes into account the spectra of the signals, as well as a fluctuating gyroscope offset,

and thereby improves the estimation accuracy. Zhao et al. [37] proposed a Kalman/UFIR filtering method for state estimation with uncertain parameters and noise statistics. Briese et al. [38] presented an adapting covariance Kalman filter based on the fusion of Ultra-Wideband (UWB) and inertial measurements. Since both sensor results are separately used in the Kalman filter, no registration between the implemented sensors is needed. Kim et al. [39] proposed a fusion algorithm based on a particle filter using vertical and road intensity information for robust vehicle localization in a large-scale urban area. However, filtering methods lack in terms of dynamic behavior and the algorithm performance varies with the change of state matrixes [38,39]. They can somewhat slow down the error accumulation process, but not eliminate it completely [40,41]. With the booming of artificial intelligence, deep neural networks [42] have been applied in the fusion of multi-sensors, while the requirement of a large scale of data still remains a big challenge in practical applications.

2.3. Cramér–Rao Lower Bound

The location of a target node is uncertain due to the influence of random factors, such as noise, fading, multi-path, and non-line-of-sight propagation, which ultimately affects the positioning accuracy [1,8]. Cramér–Rao lower bound (CRLB) defines the theoretical lower bound of the unbiased estimator variance and is used as a general criterion and benchmark for evaluating the performance of a positioning system [8].

Tichavsky et al. [43] provided the formulation of recursive posterior CRLB for nonlinear filters based on the Bayesian framework. For range-based wireless localization system, many research studies have provided CRLB results for different scenarios. Qi et al. [44] proposed a generalized CRLB (G-CRLB) of the wireless system for non-line-of-sight (NLOS) environment. Other similar works also give CRLB for different ranging techniques [45]. Although some other methods can be used for performance analysis [14], CRLB is still popular for wireless localization research due to its simplicity and general expression.

However, CRLB only focuses on the influence of the relationship between relative positions in spatial state on the accuracy of the positioning target, neglecting the time information, thus cannot meet the requirements of the time evaluation in the positioning system. The posterior Cramér–Rao lower bound (PCRLB) considers the time domain information [46] and can be used as another criterion for the performance evaluation of the positioning system. PCRLB has recently been used as the basis for determining optimal observer trajectories in bearings-only tracking (e.g., [47]). Recent interest in the PCRLB is primarily a result of an excellent paper [43] in which a computationally efficient (and general) formulation of the bound is derived. This has led to a number of further developments, including derivations of the bound and associated information reduction factors in cluttered environments [14] and the establishment of PCRLBs in multi-target tracking with either preprocessed [1,5] or raw sensor data [8,15].

To better comprehend HMT problems and provide a theoretical basis for the system verification and algorithm design, in this paper, we conduct a comprehensive analysis on performance evaluation of HMT system based on IMU/TOA fusion. By the derivation of CRLB and PCRLB, we present a feasible assessment means to evaluate IMU/TOA fusion system in both temporal and spatial aspects.

3. Problem Definition

The body can be viewed as an interconnected whole, and each movement is done in collaboration with a number of body parts. The part of body that is connected to each other is called the joint. We call this chain of rigid limbs and joints the “chain of motion”. Therefore, HMT can be considered as a collaborative tracking process of the motion chain composed of multiple human joints.

3.1. Model Description

Generally speaking, a 3D human motion is considered as a set of multiple joints and limbs movement, which can be described using mathematical expressions of coordinates and direction.

As shown in Figure 1, the whole body is divided into five connected parts, namely left upper limb, right upper limb, trunk, left lower limb and right lower limb. We use dots to represent joints and lines for limbs. The whole human body acts as a unity of these limbs and joints.

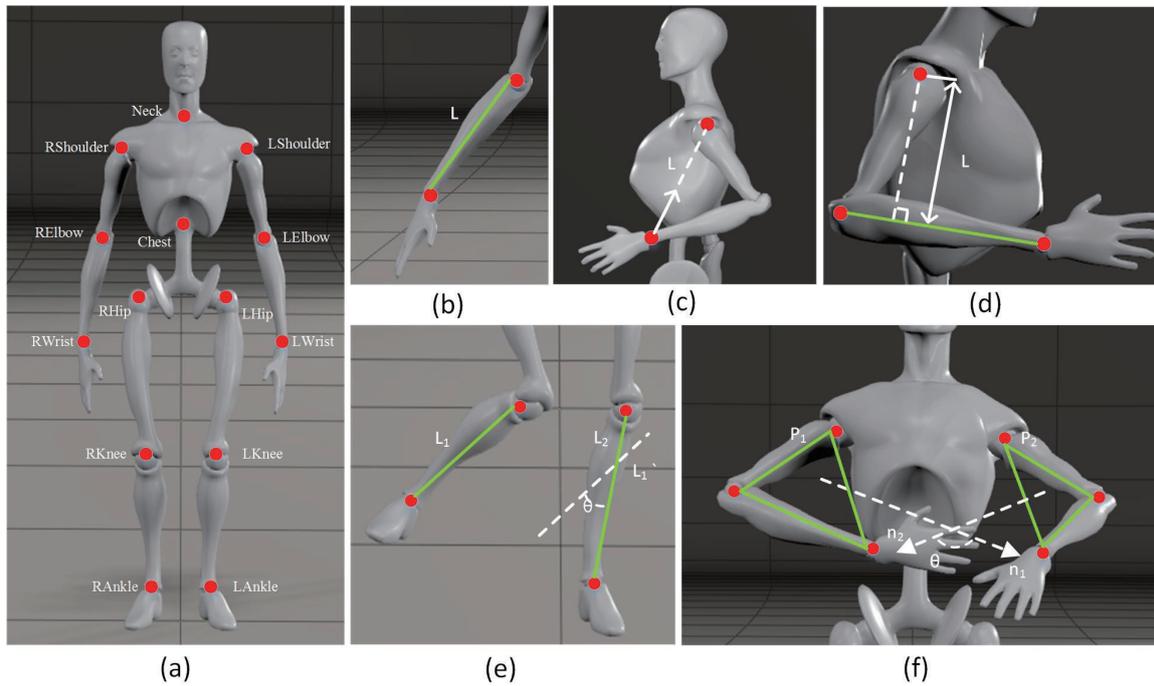


Figure 1. Human skeleton model, and the key distance and angle parameters related with human motion tracking. (a) the entire human body and (b–f) the distances and angles between joint and limbs.

As shown in Figure 1, the human body consists of 15 joints and 14 limbs. To better describe the complicated human body movements, Denavit–Hartenbe (D-H) equation [22,25] is generally used in the description of HMT features and parameters. Its measurement parameters mainly include the distances between joints (Figure 1b,c) joint and limbs (Figure 1d) and the relative angles between limbs (Figure 1e,f). To simplify the above tracking problem, we abstract the HMT serialization process as shown in Figure 1, representing the target’s coordinate and measurement parameters in time sequence $T = \{t_0, t_1, \dots, t_n\}$.

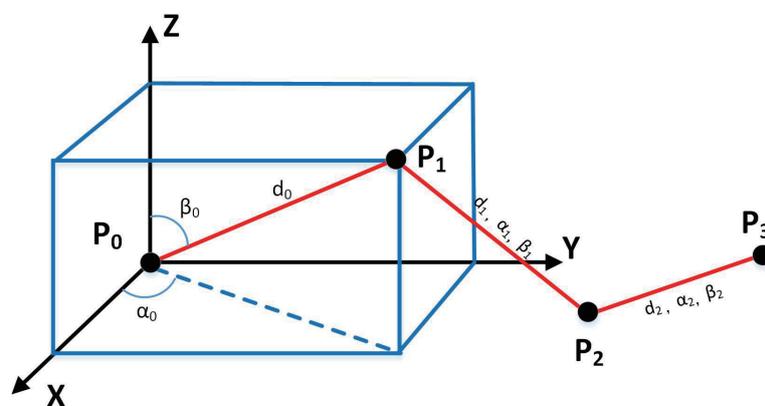


Figure 2. The definition of Gauss–Markov random process towards human motion tracking and its key parameters.

On the basis of random motion model [25], we assume a Gauss–Markov process to describe the movement of the joints of the body, as shown in Figure 2. Thanks to Inertial Measurement Unit (IMU), acceleration and angular data can be obtained and, furthermore, the rotation angle of the target position within the history frames can be calculated by means of Attitude and Heading Reference System (AHRS) [20]. In addition, distance-related information can be obtained using the TOA technique [19], after a few simple calculations. Compared with IMU-based dead reckoning method [25], TOA-based tracking method does not have to face the problem of drifting.

3.2. Error Sources

As mentioned above, the measurement parameters of the system are derived from IMU and TOA. The measurement error sources can be divided into two aspects: distance measurement error and angle measurement error. Thereinto, target node’s distance parameter \hat{d}_t at time t is defined as:

$$\hat{d}_t = d_t + n_d \tag{1}$$

where d_t is the actual distance value and n_d is Gauss distributed measurement noise whose variance is σ_d^2 , namely $n_d \sim \mathcal{N}(0, \sigma_d^2)$. Vector $\hat{d}_t = [\hat{d}_1, \hat{d}_2, \dots, \hat{d}_{n+m-1}]^T$ is introduced to collect distance information, namely the distance between target node and reference nodes, respectively, standing for one body joint.

Horizontal heading estimates from inertial-sensor-based approach α reads:

$$\hat{\alpha}_t = \alpha_t + u_t, u_t \sim N(0, \epsilon_t^2) \tag{2}$$

where α_t is the true value of actual horizontal heading. u_t is uncorrelated zero-mean Gaussian random variable with variance ϵ_t^2 , which is independent from z direction and is also uncorrelated with \hat{d} . The horizontal heading is an instantaneous measurement parameters, which may introduce cumulative error by the characteristics of IMU. Namely, $\hat{\alpha}_t$ is a time-dependent parameter. Vector $\alpha_t = [\alpha_1, \alpha_2, \dots, \alpha_{N_\alpha}]^T$ is introduced to collect α_t , i.e., the angular cumulative information in the past N_α frames.

Vertical elevation estimates from inertial-sensor-based approach β reads:

$$\hat{\beta}_t = \beta_t + v_t, v_t \sim N(0, \zeta_t^2) \tag{3}$$

where β_k is the actual vertical elevation. v_t is uncorrelated zero-mean Gaussian random variable with variances ζ_t^2 , which is independent from z direction and is also uncorrelated with both $\hat{\alpha}$ and \hat{d} . $\hat{\beta}_t$ is also a time-independent parameter. Vector $\beta_t = [\beta_1, \beta_2, \dots, \beta_{N_\beta}]^T$ is introduced to collect β_k , i.e., the angular cumulative information in the past N_β frames. As above mentioned angles α and β are all measured from IMU, thus, for ease of expression, we take σ_{IMU} as the variance of both these two parameters.

3.3. Model Presentation

In continuous human motion tracking process, some body parts remain relatively steady in comparison to the end joints or limbs of the human body, such as the trunk. Thus, these positions may be set as reference nodes, defined as $\mathcal{P}_a = \{1, \dots, m\}$, in which the k th reference node’s coordinate is defined as a_k . m indicates the number of reference nodes. In addition, the position state of the target is denoted by $p_t = [p_t^X \ p_t^Y \ p_t^Z]^T$, where p_t^X , p_t^Y and p_t^Z are the coordinates in the 3-D positioning system, and T is the transpose operator. The target nodes set is represented as $\mathcal{P}_x = \{1, 2, \dots, n\}$, and the i th target node’s coordinate is defined as p_i . n indicates the number of target nodes.

According to the Bayesian theorem [22,25], at time t , the state estimation x_{t-1} could be represented as:

$$p_t = f_t(p_{t-1}) + q_t \tag{4}$$

where Equation (4) is the prediction equation, and $f_t(\cdot)$ is the transition equation, and the noise of x_t follows standard normal distribution with variance Q_t , namely $\mathbf{q}_t \sim \mathcal{N}(0, Q_t)$.

To describe the measurement parameters of human motion in a more general way, we define the measurement parameters of the system as:

$$\mathbf{z}_t = \mathbf{h}_t(\mathbf{d}(\mathbf{p}_t, \mathbf{k}), \boldsymbol{\varphi}) + \mathbf{r}_t \tag{5}$$

where the measurement vector \mathbf{z}_t , whose noise follows normal distribution, i.e., $\mathbf{r}_t \sim \mathcal{N}(0, R_t)$, contains two parts, namely TOA measurement parameter (distance) and IMU measurement parameter (acceleration, angular velocity, etc.). $\mathbf{h}_t(\cdot)$ is nonlinear observation equation, which is related with actual measurement values of target \mathbf{p}_t , including $\mathbf{d}(\mathbf{p}_t, \mathbf{k})$ (the distance vector at time t), and $\boldsymbol{\varphi}$ (the angle vector at time t). $\mathbf{k} = [k_1, \dots, k_j, \dots, k_{m+n-1}]$ is distance measurement indicator, such as line-of-sight (LOS)/non-line-of-sight (NLOS), the reference node accuracy indicator, etc. $m + n - 1$ is the total number of both reference and target nodes. $\boldsymbol{\varphi} = [\varphi_1, \dots, \varphi_i, \dots, \varphi_s]$ is the vector of angle measurement values, whose amount is s . \mathbf{k} and $\boldsymbol{\varphi}$ are auxiliary parameters that are not essential but may help increase the localization accuracy and could be calculated in real time. The effects of these parameters can be verified in subsequent simulation experiments.

4. IMU/TOA Fusion Based HMT Method

Our analysis considers the possible unknown random factors that may influence the human motion tracking. Considering the characteristics of sequential human motion tracking, we define the measuring variable at time slot t as $\boldsymbol{\theta}_t$ (for simplicity, it is abbreviated as $\boldsymbol{\theta}$), namely

$$\boldsymbol{\theta} \triangleq [\mathbf{p}_t^T \quad \mathbf{p}_{t-1}^T \quad \mathbf{k}^T \quad \boldsymbol{\varphi}^T]^T \tag{6}$$

where \mathbf{p}_t is the target position vector at the moment, and \mathbf{p}_{t-1} is historical target location information at the former time slot.

Cramér–Rao lower bound (CRLB) [22] is defined as the inverse of Fisher Information Matrix (FIM) and represents the theoretical lower bound of the observed variance of unbiased estimators [22,25]. Suppose that $p(\boldsymbol{\theta}, \mathbf{z}_t)$ represents the joint probability density function of observation vectors \mathbf{z}_t and $\boldsymbol{\theta}$, then FIM could be derived as the variance of its log likelihood function gradient, namely

$$\mathbf{J}(\boldsymbol{\theta}) \triangleq \mathbb{E}[\Delta_{\boldsymbol{\theta}}^{\theta} \ln p(\boldsymbol{\theta}, \mathbf{z}_t)] \tag{7}$$

where $\mathbb{E}\{\cdot\}$ represent the expectation of parameters. $\nabla_{\boldsymbol{\theta}} = [\frac{\partial}{\partial \theta_1}, \dots, \frac{\partial}{\partial \theta_n}]$, $\Delta_{\alpha}^{\beta} = \nabla_{\beta} \nabla_{\alpha}^T$. Thus,

$$\mathbf{Cov}_{\boldsymbol{\theta}}(\tilde{\boldsymbol{\theta}}) \succeq \{\mathbf{J}(\boldsymbol{\theta})\}^{-1} \tag{8}$$

where $A \geq B$ for the sake of simplicity, represents that matrix $A - B$ is non-negatively defined.

According to the Bayesian theorem, $p(\boldsymbol{\theta}, \mathbf{z}_t) = p(\mathbf{z}_t|\boldsymbol{\theta})p(\boldsymbol{\theta})$, then $\mathbf{J}(\boldsymbol{\theta})$ could be divided into two parts, namely

$$\mathbf{J}(\boldsymbol{\theta}) = \mathbf{J}_D(\boldsymbol{\theta}) + \mathbf{J}_P(\boldsymbol{\theta}) \tag{9}$$

where \mathbf{J}_D is the information matrix obtained by measuring parameters. \mathbf{J}_P is the information matrix obtained by prior information. We describe the representation of these two variables separately in the following section.

4.1. Measurement Parameter Information Matrix

Due to the measurement equation, we get $\mathbf{h} = \mathbf{h}_t(\mathbf{d}(\mathbf{p}_t, \mathbf{k}), \boldsymbol{\varphi})$, and then, based on the chain rules [48], \mathbf{J}_D could be represented as:

$$\mathbf{J}_D = \mathbf{H} \cdot \mathbf{J}_h \cdot \mathbf{H}^T \tag{10}$$

where $\mathbf{H} = [\nabla_{\theta} h]$, J_h is the FIM of h , namely

$$\mathbf{J}_h = \mathbb{E} \left\{ \nabla_{\mathbf{h}} \ln p(\mathbf{z}_t | \boldsymbol{\theta}) [\nabla_{\mathbf{h}} \ln p(\mathbf{z}_t | \boldsymbol{\theta})]^T \right\} \tag{11}$$

where transition matrix \mathbf{H} could be further decomposed into four components

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_t & \mathbf{H}_{t-1} & \mathbf{K} & \Phi \end{bmatrix}^T \tag{12}$$

where $\mathbf{H}_t = [\nabla_{p_t} h]$, $\mathbf{H}_{t-1} = [\nabla_{p_{t-1}} h]$, $\mathbf{K} = [\nabla_k h]$, $\Phi = [\nabla_{\varphi} h]$. Because the distance is a instantaneous measurement parameter, it has no error accumulation problems. Thus, \mathbf{d} has no relationship with prior information p_{t-1} , namely $\mathbf{H}_{t-1} = 0$. Then,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_t & \mathbf{0} & \mathbf{K} & \Phi \end{bmatrix}^T \tag{13}$$

For J_h , we can use diagonal matrices of order n to represent it,

$$\mathbf{J}_h = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_j, \dots, \lambda_n) \tag{14}$$

where λ_j stands for the j th measurement value. Then, \mathbf{J}_D is represented as a $(4 + m + n + s) \times (4 + m + n + s)$ matrix,

$$\begin{aligned} \mathbf{J}_D &= \mathbf{H} \cdot \boldsymbol{\Lambda} \cdot \mathbf{H}^T \\ &= \begin{bmatrix} \mathbf{D}_{11} & \mathbf{0} & \mathbf{D}_{13} & \mathbf{D}_{14} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_{13}^T & \mathbf{0} & \mathbf{D}_{33} & \mathbf{D}_{34} \\ \mathbf{D}_{14}^T & \mathbf{0} & \mathbf{D}_{34}^T & \mathbf{D}_{44} \end{bmatrix} \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathbf{D}_{11} &= \mathbf{H}_t \boldsymbol{\Lambda} \mathbf{H}_t^T & \mathbf{D}_{13} &= \mathbf{H}_t \boldsymbol{\Lambda} \mathbf{K}^T \\ \mathbf{D}_{14} &= \mathbf{H}_t \boldsymbol{\Lambda} \Phi^T & \mathbf{D}_{33} &= \mathbf{K} \boldsymbol{\Lambda} \mathbf{K}^T \\ \mathbf{D}_{34} &= \mathbf{K} \boldsymbol{\Lambda} \Phi^T & \mathbf{D}_{44} &= \Phi \boldsymbol{\Lambda} \Phi^T \end{aligned} \tag{16}$$

4.2. State Parameter Information Matrix

It could be known by the whole probability formula, for the variable $\boldsymbol{\theta}$, the probability $p(\boldsymbol{\theta}) = p(\mathbf{p}_t | \mathbf{p}_{t-1}) p(\mathbf{k}) p(\boldsymbol{\varphi})$. Then, log-likelihood function could be represented as:

$$\ln p(\boldsymbol{\theta}) = [\ln p(\mathbf{p}_t | \mathbf{p}_{t-1})] + \ln p(\mathbf{k}) + \ln p(\boldsymbol{\varphi}) \tag{17}$$

where $p(\mathbf{k})$ and $p(\boldsymbol{\varphi})$ are corresponding independent information of \mathbf{p}_t and \mathbf{p}_{t-1} . We decompose the vector $\boldsymbol{\theta}$ into two components, namely state vector $[\mathbf{p}_t \ \mathbf{p}_{t-1}]^T$ and observation vector $[\mathbf{k} \ \boldsymbol{\varphi}]^T$. Thus, J_P could be represented as:

$$\begin{aligned} \mathbf{J}_P &= \mathbb{E}[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{\theta})] \\ &= \begin{bmatrix} \mathbf{J}_{P11} & \mathbf{J}_{P12} \\ \mathbf{J}_{P12}^T & \mathbf{J}_{P22} \end{bmatrix} \end{aligned} \tag{18}$$

where J_{P11} could be obtained with the recursion of vectors \mathbf{p}_t and \mathbf{p}_{t-1} , which, according to Tichavsky et al. [43], could be described as:

$$\mathbf{J}_{P11} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{12}^T & \mathbf{M}_{22} + \mathbf{J}(\mathbf{p}_{t-1}) \end{bmatrix} \tag{19}$$

where

$$\begin{aligned} \mathbf{M}_{11} &= \mathbf{Q}_t^{-1} \\ \mathbf{M}_{12} &= \nabla_{\mathbf{p}_{t-1}} \mathbf{f}_t(\mathbf{p}_{t-1}) \mathbf{Q}_t^{-1} \\ \mathbf{M}_{22} &= \nabla_{\mathbf{p}_{t-1}} \mathbf{f}_t(\mathbf{p}_{t-1}) \mathbf{Q}_t^{-1} [\nabla_{\mathbf{p}_{t-1}} \mathbf{f}_t(\mathbf{p}_{t-1})]^T \end{aligned} \tag{20}$$

$\mathbf{J}(\mathbf{p}_{t-1})$ is the FIM of \mathbf{p}_{t-1} . As $p(k)$ and $p(\varphi)$ are independent from \mathbf{p}_t and \mathbf{p}_{t-1} , $\mathbf{J}_{P_{12}}$ is $\mathbf{0}$, namely

$$\mathbf{J}_{P_{12}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{21}$$

Similarly, the prior probability $p(\mathbf{p}_t | \mathbf{p}_{t-1})$ are independent from k and φ . Thus, $\mathbf{J}_{P_{12}} = \mathbf{J}_{P_{21}}^T = \mathbf{0}$. Finally,

$$\mathbf{J}_{P_{22}} = \begin{bmatrix} \mathbf{J}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_\varphi \end{bmatrix} \tag{22}$$

where \mathbf{J}_K and \mathbf{J}_φ are, respectively, the FIM of vectors k and φ :

$$\begin{aligned} \mathbf{J}_K &= \mathbb{E}[\Delta_k^k \ln p(k)] \\ \mathbf{J}_\varphi &= \mathbb{E}[\Delta_\varphi^\varphi \ln p(\varphi)] \end{aligned} \tag{23}$$

4.3. Integral Information Matrix

Due to the analysis above, substituting Equations (15) and (18) into Equation (9), the FIM of variable θ could be represented as:

$$\mathbf{J}(\theta) = \begin{bmatrix} \mathbf{M}_{11} + \mathbf{D}_{11} & \mathbf{M}_{12} & \mathbf{D}_{13} & \mathbf{D}_{14} \\ \mathbf{M}_{12}^T & \mathbf{M}_{22} + \mathbf{J}(x_{t-1}) & \mathbf{0} & \mathbf{0} \\ \mathbf{D}_{13}^T & \mathbf{0} & \mathbf{D}_{33} + \mathbf{J}_K & \mathbf{D}_{34} \\ \mathbf{D}_{14}^T & \mathbf{0} & \mathbf{D}_{34}^T & \mathbf{D}_{44} + \mathbf{J}_\varphi \end{bmatrix}. \tag{24}$$

However, only the lower bound of \mathbf{p}_t is of interest, namely a small submatrix of $\mathbf{J}(\theta)$, $\mathbf{J}^{-1}(\mathbf{p}_t) = [\mathbf{J}^{-1}(\theta)]_{2 \times 2}$.

Furthermore, due to the Schur complement theorem [49], $\mathbf{J}(\mathbf{p}_t)$ could be divided into two components, namely state matrix \mathbf{J}_S and measurement matrix \mathbf{J}_C , then

$$\mathbf{J}(\mathbf{p}_t) = \mathbf{J}_S - \mathbf{J}_C \tag{25}$$

where

$$\begin{aligned} \mathbf{J}_S &= \mathbf{M}_{11} + \mathbf{D}_{11} - \mathbf{M}_{12} (\mathbf{M}_{22} + \mathbf{J}(\mathbf{p}_{t-1}))^{-1} \mathbf{M}_{12}^T \\ \mathbf{J}_A &= \begin{bmatrix} \mathbf{D}_{13} & \mathbf{D}_{14} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{33} + \mathbf{J}_K & \mathbf{D}_{34} \\ \mathbf{D}_{34}^T & \mathbf{D}_{44} + \mathbf{J}_\varphi \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{D}_{13} & \mathbf{D}_{14} \end{bmatrix}^T \end{aligned} \tag{26}$$

As mentioned above, we demonstrate a comprehensive analysis of how distance and angle measurements are related with human motion tracking accuracy in IMU/TOA fusion systems. The detailed calculation process is shown in Algorithm 1. In the next two sections, we further evaluate the overall performance of the fusion system from two aspects, namely spatial and temporal performance, in theory.

Algorithm 1 Proposed IMU/TOA fusion-based human motion tracking method.

0: **Initialization:**

1: Define measuring variable at time slot t as $\theta_t \triangleq \begin{bmatrix} p_t^T & p_{t-1}^T & k^T & \phi^T \end{bmatrix}^T$. Set the time slot index $t = 0$. Choose an initial value for p_{t-1} of variable θ_t , generally as 0.

Part 1: measurement parameter information

2: Calculate parameter matrix H_t, K, Φ .

3: Update the matrix parameters $H = \begin{bmatrix} H_t & H_{t-1} & K & \Phi \end{bmatrix}^T$.

4: Calculate parameter matrix $D_{11}, D_{13}, D_{14}, D_{33}, D_{34}$ and D_{44} .

5: Update the matrix parameter J_h .

6: Update the information matrix obtained by measuring parameters, i.e., $J_D(\theta) = H \cdot J_h \cdot H^T$

Part 2: state parameter information

7: Calculate parameter matrix $M_{11}, M_{12}, M_{22}, J_K$ and J_Φ .

8: Update the matrix parameters $J_{P_{11}}, J_{P_{12}}$ and $J_{P_{12}}$.

9: Update the information matrix obtained by prior information, i.e., $J_P = \begin{bmatrix} J_{P_{11}} & J_{P_{12}} \\ J_{P_{12}}^T & J_{P_{22}} \end{bmatrix}$.

Part 3: integral information

10: Update the FIM matrix by $J(\theta) = J_D(\theta) + J_P(\theta)$

11: Update Schur complement parameters of $J(\theta)$, namely state matrix J_S and measurement matrix J_C

12: Update the estimation result $p_t = J(p_t) = J_S - J_C$

Part 4: recursive computation

13: $t + 1 \rightarrow t$, go to step 2.

5. Spatial Performance Analysis

When it comes to spatial performance, we suppose that the tracking system has time independence, namely $J(p_t)$ and $J(p_{t-1})$ are independent from each other. Thus, $M_{12}, M_{22}, J_{p_{t-1}}, J_K, D_{13}, D_{33}$ and D_{34} are all 0, namely

$$J(p_t) = M_{11} + D_{11} - D_{14}[D_{44} + J_\Phi]^{-1}D_{14}^T \quad (27)$$

Movement characteristics of each body part should be taken into consideration. Because of the spatial features of human motion sensing, it could be classified into two categories: 2-D and 3-D conditions. 2D motions refer to the body movements that can be captured by a plane, such as raising one's arms into Y pose or T pose. 3D motions refer to stereoscopic movements such as walking, running or swing hands. The different relative positions may contribute to various capture accuracy. Therefore, both the CRLB of human body in 2D and 3D scenarios are taken into consideration. CRLB of each location is derived as following demonstration and performance of proposed fusion method is verified. See Appendix A for the detailed derivation process of CRLB for spatial performance analysis.

5.1. Scenario Setup

As mentioned above, 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 constrained scenario of 2 m \times 2 m is set up for further experiments. The entire human body is supposed to located in the center of this region. In contrast, two sequences of reference nodes combination are considered, namely $Case1 = \{Neck, Chest, LHip, RHip\}$ and $Case2 = \{Neck, LShoulder, RShoulder, Chest, LHip, RHip\}$. The body is allowed to move freely in this space.

5.2. Performance Analysis

To demonstrate the typical calculation results, CRLB, when the variance of distance measurement σ_d is 0.2 and σ_{IMU} is 0.1, is selected upon common commercial motion sensing systems, such as Xsens. From the experiment results, as shown in Figure 3, the following conclusions could be obtained:

1. In the aspect of two-dimensional condition, shown in Figure 3a,b, lower CRLB is shown when the position is closer to the trunk. It is also lower in the lower limb than the upper, which is possibly due to the selection of reference nodes. To verify this, a comparison experiment was conducted when the reference nodes were chosen differently. Results shown in Figure 3a,b indicate that the relative position of the reference nodes cause different CRLB of human motion and the more uniform the nodes distribute, the relative lower CRLB it achieves. The same result can also be seen in Figure 3c,d when only distance measurement is applied in the human motion sensing process.
2. Since human motion is a three-dimensional process, stereoscopic presentation is shown in Figure 3e,f. The human is assumed to be placed in the XOZ plane when the y is set as zero. The 3D version of CRLB is likely to be a 2D one that stretches along the Z-axis. Similar results could be seen in 3D condition that, the closer to the trunk, the lower the CRLB that could be achieved.
3. For comparison, the CRLB under two measurement method were calculated, respectively, independent distance measurement and the fusion of IMU and distance measurement. Comparison results are shown in Figure 3a,c (also in pair of Figure 3b,d).

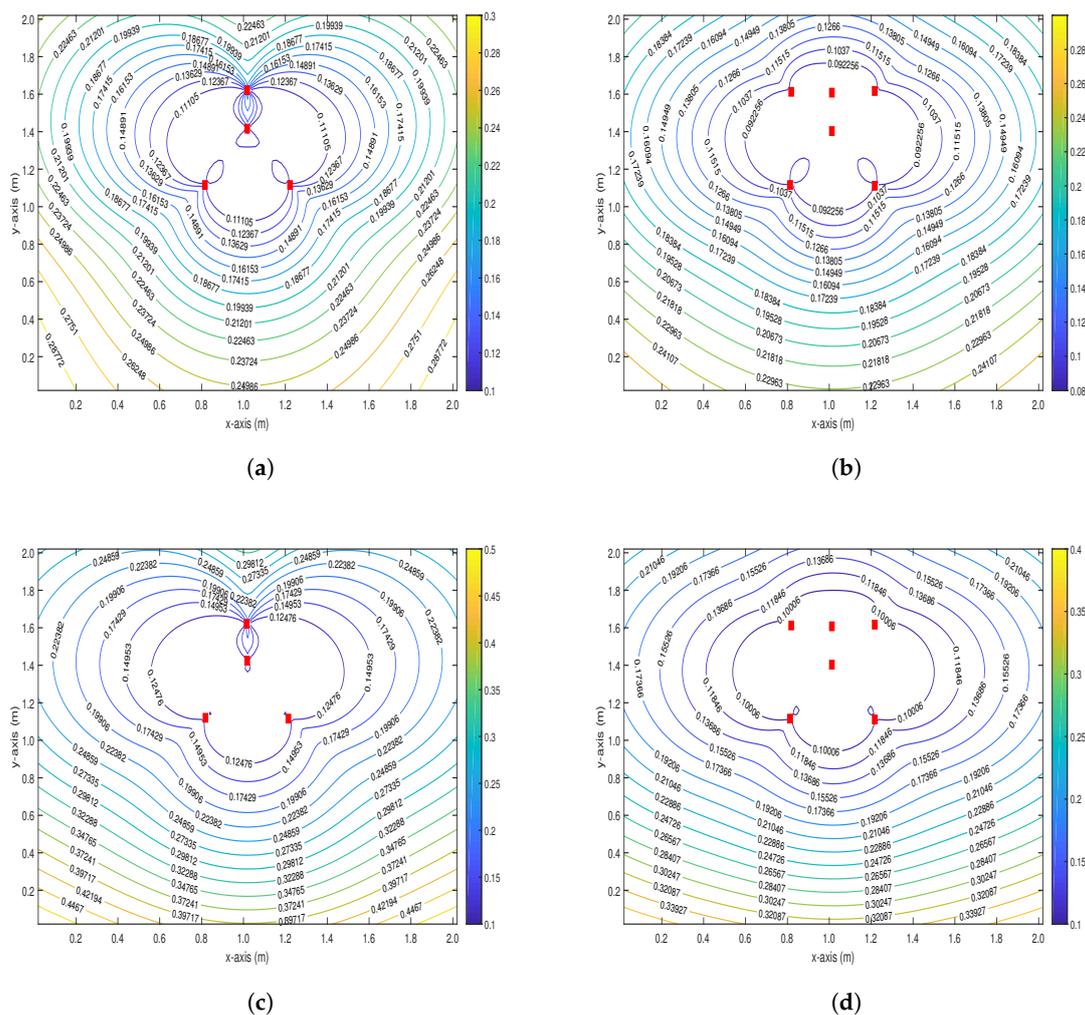


Figure 3. Cont.

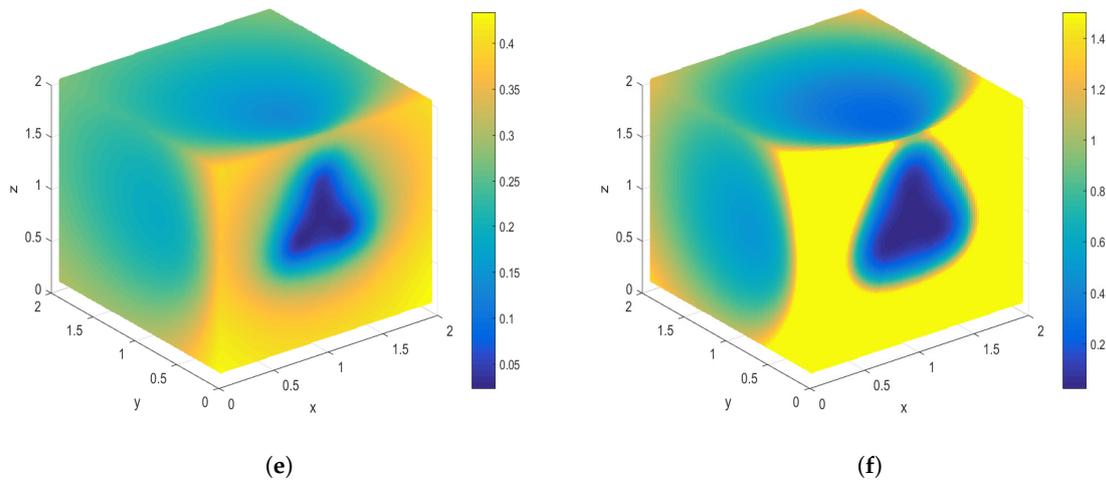


Figure 3. CRLB comparison in two-dimensional and three-dimensional condition: (a,b,e) CRLB in condition that both distance and IMU information are considered using proposed model; and (c,d,f) CRLB considered with distance information solely. σ_d is set as 0.2 and σ_{IMU} is 0.1.

To clearly show the different capture capability of distance measurement and IMU, Figure 4 is presented. As shown, under different σ_d^2 , CRLB increases along with the increase of σ_{IMU}^2 . When σ_{IMU}^2 is fixed, CRLB also increases along with the increase of σ_d^2 . However, the change is more apparent when σ_d^2 is fixed. It means that the accuracy of IMU measurement matters more in the fusion process, but the distance measurement helps restrain the bounds of CRLB. It is worth mentioning that, for a given σ_{IMU}^2 , CRLB shows little variation with the increase of σ_d^2 if it is larger than 0.2. It indicates that the distance measurement may be not the main limiting factor of CRLB, but σ_{IMU}^2 is.

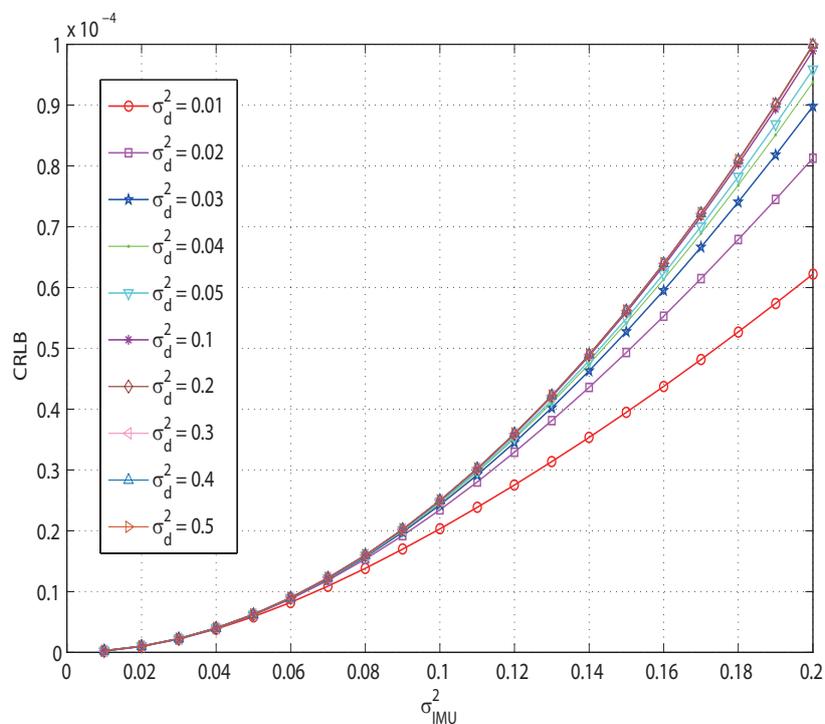


Figure 4. CRLB distribution under different σ_d^2 when σ_{IMU}^2 changes.

6. Temporal Performance Analysis

We suppose that the tracking system has cumulative recursion characteristics, namely $J(\mathbf{p}_t)$ and $J(\mathbf{p}_{t-1})$ are not independent from each other. Based on [22], Equation (26) could be simplified into a recursive form, namely Posterior Cramér–Rao Lower Bound (PCRLB) [22,25].

$$\mathbf{J}(\mathbf{x}_t) = \mathbf{M}_{11} + \mathbf{D}_{11} - \mathbf{M}_{12} [\mathbf{M}_{22} + \mathbf{J}(\mathbf{x}_{t-1})]^{-1} \mathbf{M}_{12}^T \quad (28)$$

See Appendix B for the detailed derivation process of PCRLB for temporal performance analysis.

6.1. Scenario Setup

Even though there are elegant expressions to recursively calculate the FIM, Equation (27) usually does not have analytical close-form solution. To deal with that, we employ the Monte Carlo approach [50] to convert those continuous integrals into discrete summations, and finally work out the PCRLB. The root-mean-square of PCRLB is given by $\frac{1}{L} \sum_{i=1}^L P_k^i$, where P_k^i is the PCRLB on the Root-Mean-Square Error (RMSE) of joint at step k in the i th Monte Carlo trial. L is the total Monte Carlo trial number ($L = 1000$ was used in this study). Note that, for each Monte Carlo trial, we randomly selected the initial location for PCRLB calculation to get a fair average of the entire human body moving process.

6.2. Performance Analysis

We mainly analyzed the influence of different factors on PCRLB from the following two perspectives: fusion method factor and topology factor.

(1) Fusion method factor: Considering different Q_t and R_t in IMU/TOA fusion methods, as well as using sole TOA or IMU method, may lead to different results of human motion tracking accuracy.

Figure 5 shows in IMU/TOA fusion methods, the lower bounds under different Q_t and R_t as well as in the use of sole TOA or IMU method. As shown in the figure:

1. When only inertial sensors are adopted in the tracking system, as indicated by the black solid line in Figure 5, the accumulative errors may tend to be diverging. Theoretically, this confirms that IMU based HMT system faces the problem of accumulative errors. However, IMU/TOA fusion method can avoid this divergence. The performance curves of proposed approaches achieve stability after certain steps, i.e., their errors converge.

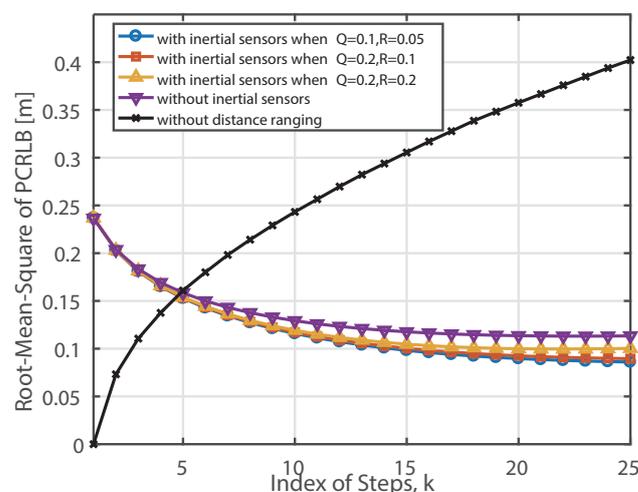


Figure 5. Temporal performance as a function of step index.

2. Compared with sole TOA tracking method, IMU/TOA fusion-based method can significantly increase the accuracy of human body motion tracking. The lower bound of proposed fusion method could drop below 8 cm.
3. With increase of Q_t and R_t , PCRLB also increases, which means that the accuracy is inversely proportional to these two parameters.

(2) Topology factor: As already indicated in Section IV-B, topologies have much influence on the performance of certain tracking systems. We work out the PCRLB of different topologies shown in Figure 6, where Q_t is 0.1 and R_t is 0.05. As shown in Figure 7, we could draw the following conclusions:

1. Under different topologies, the theoretical accuracy lower bounds of the tracking system are different.
2. The different topological conditions tend to be stable after the same iteration number (around 15), which means the convergence rate is roughly the same under various topologies.
3. Topology 3 suffers the largest RMSE, which may be due to the dense reference nodes. The minimum RMSE is less than 8 cm, as shown in Topology 4, which makes it a better choice for practice applications.
4. The proposed fusion HMT systems with TOA integrated share identical trend results with that of solely TOA ranging.

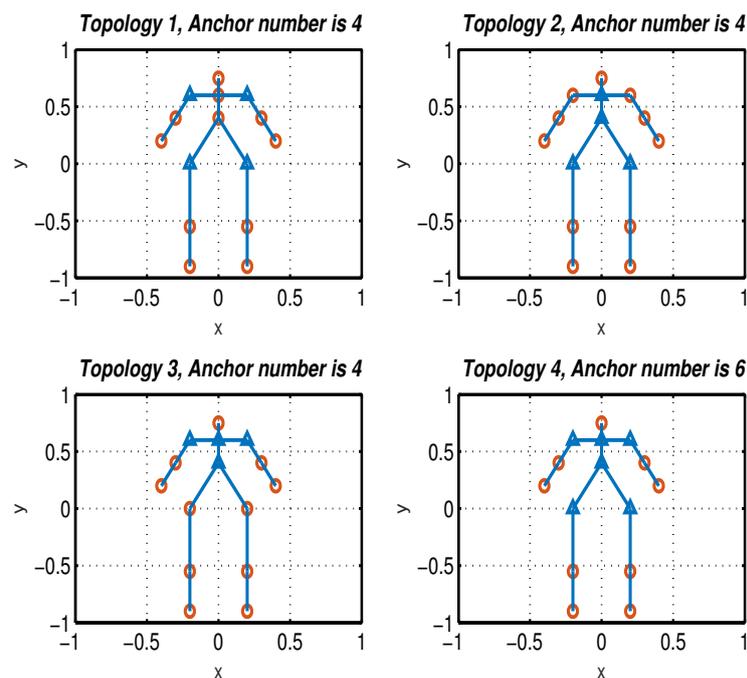


Figure 6. Topologies configuration. \circ represents free joints, and \triangle represents selected anchors.

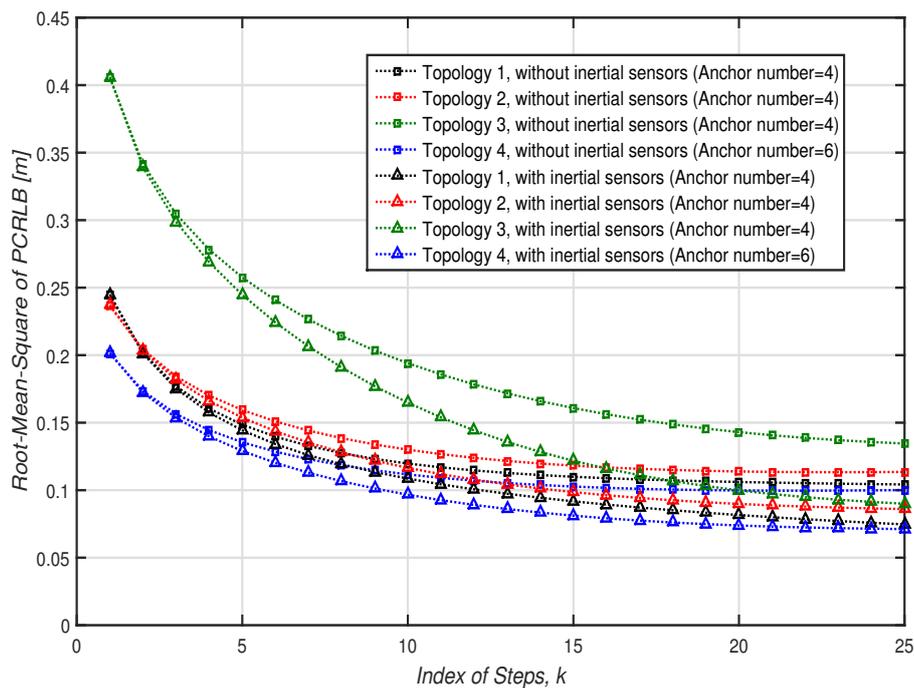


Figure 7. Root mean square of PCRLB as a function of index of steps. Different receiver topologies are considered.

7. Practical Use Case

In above sections, the theoretical effectiveness of our proposed human motion tracking method is verified in terms of both spatial and temporal performances. The introduced IMU/TOA fusion method utilizes both distance and IMU information to lower the expected error bounds in theory. In this section, we take advantage of this method in 3D practical application scenarios to verify its performance.

7.1. Platform Overview and Experiment Setup

A minimized wearable sensing platform was specially designed for the tracking of human motion in 3D scenario. It aimed to collect spatial and temporal information during the human motion process. Sensing nodes were designed and intended to be put on joints to capture the movement conditions. Target data included accelerated velocity, angular velocity and the distances between joints and body parts. Among these, distance information was especially different when compared with other platforms, such as Xsens and Invensense. We sampled data using above-mentioned platform at 10 Hz. All processing work were performed online in Matlab with PC (Intel Core i7-6700M CPU, 16GB RAM). The communication between wearable nodes and the PC was via Bluetooth.

Our integrated wearable sensor system was composed of two parts: one control unit and several data acquisition units. The control unit worked as a gateway to control the whole operation process via Bluetooth communication. Data acquisition unit was mainly responsible for data sensing. Each data acquisition unit had a six-axis sensor (MPU6050, which integrated a triaxial accelerometer and a triaxial gyroscope), a barometer sensor (MS5611) and a UWB-TOA ranging module (DWM1000) [19]. The MEMS sensors were connected to a micro-controller (STM32F103) for the sake of sampling efficiency at a rate of 10 Hz. Data were transferred to the control unit in real-time and also written into its SD card as backups for other offline analysis and applications. Three sensing nodes were mounted onto the ankle, knee and hip joint. Distance measurements were conducted between these nodes. The whole system architecture is demonstrated in Figure 8.

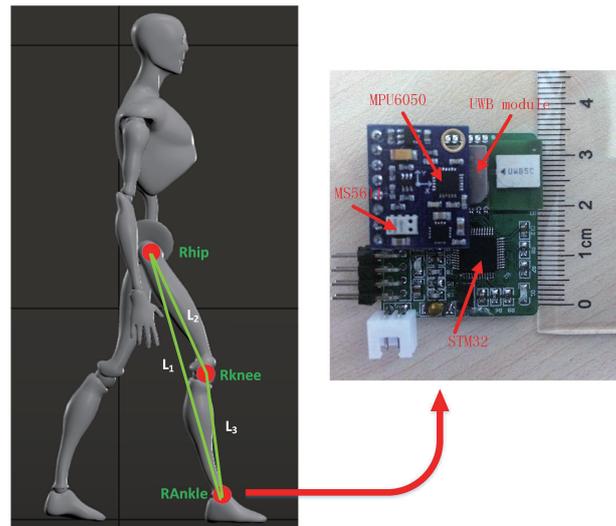


Figure 8. Experimental Platform Settings. Three sensing nodes were mounted onto the ankle, knee and hip joint. Each node mainly consisted of micro-controller unit (STM32), IMU unit (MPU6050 and MS5611) and TOA unit (UWB module).

We proposed an IMU/TOA fusion method for human motion tracking applications, introducing distance information into inertial systems. However, one of the most obvious difference between the proposed method and that in [23] is: the algorithm in [23] needs to take advantage of external UWB stations as reference nodes, namely anchors. However, in our method, we simply utilized internal tags (mounted onto body joints) to range and conduct the fusion process, which can significantly reduce the complexity of tracking systems.

Based on above considerations, human lower limb movement analysis in 3D scenario was chosen as a typical verification experiment, for better demonstrate the performance of proposed method when it is applied to human motion tracking.

7.2. Practical Use Case in 3D Scenario

A spiral-stairs (between two floors) scenario Was selected as a 3D use case and four UWB anchors were located at each floor (totally eight anchors were used), as shown in Figure 9. The location of anchors in each floor are referred to those in [23] and only used in comparative experiments. Three comparative algorithms were also tested:

- (1) Only IMU was used for human lower limb tracking. Zero velocity update (ZUPT) algorithm [51] was applied. For simplicity, we denoted it as “IMU” method.
- (2) IMU and TOA fusion method in [23] was applied. UWB based TOA tag nodes were mounted to the right lower limb and communicated with external anchor nodes, implementing TOA localization algorithm. For simplicity, and to separate it from our proposed method, we denoted it as “IMU/ex-TOA” method.
- (3) Optimal Enhanced Kalman Filter (OEKF)-based method in [52] was applied. The cumulative errors in attitude and velocity were corrected using the attitude fusion filtering algorithm and ZUPT, respectively. For simplicity, and to separate it from our proposed method, we denoted it as “IMU-OEKF” method.

Experiment results are shown in Figures 9 and 10.

For better comparison, ground truth and tracking trajectories applying all mentioned methods (the proposed method and three comparative ones) are drawn in Figure 9. It is clearly seen that, when applying the proposed method, the experiment results were far closer to the ground truth, while the results when applying methods of IMU [51], IMU/ex-TOA [23] and IMU-OEKF [52] were drifting

away as time accumulates. The result remained similar at the very beginning; however, the gap became larger over time.

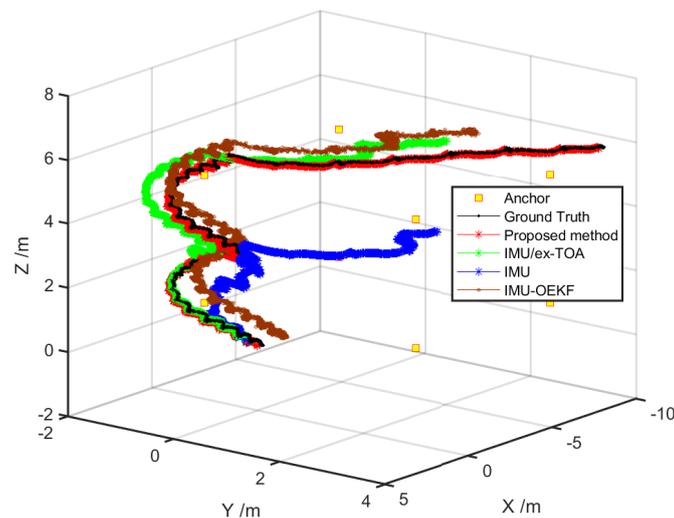


Figure 9. Walking trajectory when climbing a spiral-stairs with our proposed method and two comparative methods.

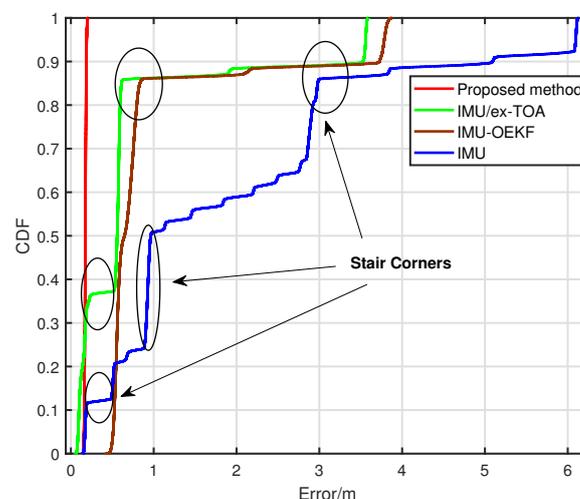


Figure 10. The accumulative distribution function (CDF) curve in 3D scenario when algorithms were applied.

Detailed accumulated errors are shown in Figure 10. IMU, IMU/ex-TOA and IMU-OEKF faced the problem of drift error, especially when the experimenter turned around at stair corners. Time accumulated errors were more likely to be caused when moving direction changes. With IMU/ex-TOA and IMU-OEKF methods applied into the algorithm, localization errors to some extent could be fixed, but still existed and were critical at turns. On the contrary, our proposed human motion tracking method showed a good stability and no clear drift errors were observed in our experiment. This is because, with the use of proposed method, body connections are considered and temporal transitions are estimated. Especially at turns, body constraints could to some extent fix time accumulated errors. The results suggest that our proposed method performed well while applied in human lower limb analysis. This also confirmed the conclusion obtained in theoretical simulation (Sections 4 and 5).

Methods with only inertial sensors faced the problem of drift errors, but our proposed model could fix time accumulative errors.

From above analysis, our proposed method had significantly higher accuracy, as well as little drift problem. Besides, compared with IMU/ex-TOA method [23] and IMU-OEKF method [52], our method did not need external anchors and had higher tracking accuracy, thus is more suitable for wearable motion tracking applications.

8. Conclusions

A IMU/TOA fusion-based human motion tracking method is proposed in this paper. IMU based HMT technique faces the tough problem of accumulative errors and drift. Time of Arrival (TOA) is considered to compensate the drift and accumulation caused by IMU. On this basis, Cramér–Rao lower bound (CRLB) is derived in detail with consideration of both temporal and spatial related factors. Simulation results show that the IMU/TOA fusion-based HMT method proposed in this paper has significantly higher accuracy, as well as little drift problem, compared with the traditional independent IMU or TOA tracking methods. The comprehensive fundamental limits analysis conducted in this paper can provide a theoretical basis for further system design and algorithm analysis. In the practical use case, proposed IMU/TOA fusion method also shows obvious performance advantages when compared with state-of-the-art methods. Besides, it does not need external anchors, thus it is more suitable for wearable motion tracking applications.

Author Contributions: C.X. and J.H. conceived and designed the experiments; C.X. and X.Z. (Xinghang Zhou) analyzed the data and contributed to the writing the manuscript. X.Z. (Xiaotong Zhang) and S.D. reviewed the analyzed data and contributed to the writing. C.X. and J.H. are co-first authors and contributed equally to this work.

Funding: This work was supported by The National Key R & D Program of China, No. 2016YFC0901303; National Natural Science Foundation of China (NSFC) project No. 61671056, No. 61302065, No. 61304257, and No. 61402033; Beijing Natural Science Foundation project No. 4152036; and Tianjin Special Program for Science and Technology No. 16ZXCXSF00150.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Derivation of the CRLB

CRLB is represented as a theoretical lower limit for any unbiased estimation and is widely used to assess localization performance. Thus, we comprehensively derive the CRLB for 3D localization of IMU/TOA fusion method in WSNs to evaluate its spatial performance. The following are some definitions. If p_k is an unbiased estimate of p_k , then

$$E(\hat{p}_k - p_k)^2 \geq CRLB = \text{tr} \left\{ J(p_k)^{-1} \right\} \quad (\text{A1})$$

where $J(p_k)$ is the Fisher information matrix [43,46]. Before solving the Fischer information matrix, we need to first define the joint probability density function as

$$p(\hat{d}_k, \hat{\alpha}_k, \hat{\beta}_k, \hat{p}_k) = \left\{ \prod_{n=1}^N p(\hat{d}_{n,k} | p_k) \right\} p(\hat{\alpha}_k | p_{k-1}, p_k) p(\hat{\beta}_k | p_{k-1}, p_k) \quad (\text{A2})$$

where $p(\hat{\alpha}_k | p_{k-1}, p_k)$, $p(\hat{\beta}_k | p_{k-1}, p_k)$ and $p(\hat{d}_{n,k} | p_k)$ can be obtained according to Equations (1)–(3).

According to the joint probability density function, we can define the Fischer information matrix as

$$J(p_k)_{i,j} = -E \left[\frac{\partial^2 \ln p(\hat{d}_k, \hat{\alpha}_k, \hat{\beta}_k, \hat{p}_k)}{\partial p_{k,i} \partial p_{k,j}} \right], i, j = 1, 2, 3 \quad (\text{A3})$$

Bringing Equation (A2) into Equation (A3) can calculate each element of the Fischer Information Matrix, and further obtain CRLB.

Appendix B. Derivation of the PCRLB

Sequential tracking problem is a temporal as well as spatial problem. This continuous information could be used to evaluate the performance of given algorithms. Thus, we extend the above CRLB to PCRLB by considering posterior information.

Before the derivation, we redefine the joint probability density function as

$$p(\hat{d}, \hat{\alpha}, \hat{\beta}, \hat{p}) = p(\hat{d}_0 | \hat{p}_0) \prod_{k=1}^K p(\hat{\alpha}_k | p_{k-1}, p_k) p(\hat{\beta}_k | p_{k-1}, p_k) p(\hat{d}_k | p_k) \tag{A4}$$

To calculate the Fischer information matrix at state k , we define

$$p_k = p(\hat{d}_{0:k}, \hat{\alpha}_{0:k}, \hat{\beta}_{0:k}, \hat{p}_{0:k}) \tag{A5}$$

where $\hat{d}_{0:k}$, $\hat{\alpha}_{0:k}$, $\hat{\beta}_{0:k}$, and $\hat{p}_{0:k}$ represent ranging, vertical angle, horizontal angle and target coordinate vector from the start state to state k , respectively.

Therefore,

$$J(p_{0:k}) = \begin{bmatrix} E \left\{ -\Delta_{p_{0:k-1}}^{p_{0:k-1}} \ln p_k \right\} & E \left\{ -\Delta_{p_{0:k-1}}^{p_{0:k}} \ln p_k \right\} \\ E \left\{ -\Delta_{p_{0:k}}^{p_{0:k-1}} \ln p_k \right\} & E \left\{ -\Delta_{p_{0:k}}^{p_{0:k}} \ln p_k \right\} \end{bmatrix} = \begin{bmatrix} A_k & B_k \\ B_k^T & C_k \end{bmatrix} \tag{A6}$$

According to Tichavsky et al. [43], the sub-matrix J_k can be obtained by pseudo-inverse of the matrix $J(p_{0:k})$, i.e.,

$$J_k = C_k - B_k^T A_k^{-1} B_k \tag{A7}$$

According to Equations (A5) and (A6), the joint probability density for the $k + 1$ state is

$$p_{k+1} = p_k p(\hat{\alpha}_{k+1} | p_k, p_{k+1}) p(\hat{\beta}_{k+1} | p_k, p_{k+1}) p(\hat{d}_{k+1} | p_k, p_{k+1}) \tag{A8}$$

According to the joint probability density of state $k + 1$, we can find that

$$J(p_{0:k+1}) = \begin{bmatrix} A_k & B_k & 0 \\ B_k^T & C_k + H_k^{11} & H_k^{12} \\ 0 & H_k^{12} & \gamma_{k+1} + H_k^{12} \end{bmatrix} \tag{A9}$$

where H_k^{11} , H_k^{12} , and H_k^{22} reflect the posterior information from state k to state $k + 1$, and γ_{k+1} reflects the location information based on TOA ranging [19].

From $J(p_{0:k+1})$ and J_k , we can get the Fischer information matrix for state $k + 1$, i.e.,

$$J_{k+1} = \gamma_{k+1} + H_k^{22} - \begin{bmatrix} 0 & H_k^{12} \end{bmatrix} \begin{bmatrix} A_k & B_k \\ B_k^T & C_k + H_k^{11} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ H_k^{12} \end{bmatrix} = \gamma_{k+1} + H_k^{22} - H_k^{12} (J_k + H_k^{11})^{-1} H_k^{12} \tag{A10}$$

Due to the step error and the directional error obey Gaussian distribution, $H_k^{11} = H_k^{12} = H_k^{22} = H_k$ can be calculated.

The solution of H_k can be found in [53]. In summary, the posterior Fisher information matrix is

$$J_{k+1} = \gamma_{k+1} + H_k - H_k(J_k + H_k)^{-1}H_k \quad (\text{A11})$$

According to the SMW (Sherman–Morrison–Woodbury) formula [54], it can be further simplified as

$$J_{k+1} = \gamma_{k+1} + (H_k^{-1} + J_k^{-1})^{-1} \quad (\text{A12})$$

where γ_{k+1} reflects the information based on TOA, H_k reflects information based on IMU.

Then, sequential simulation can be conducted by Monte Carlo methods [50].

References

1. Gravina, R.; Alinia, P.; Ghasemzadeh, H.; Fortino, G. Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. *Inf. Fusion* **2017**, *35*, 68–80. [CrossRef]
2. Fortino, G.; Giannantonio, R.; Gravina, R.; Kuryloski, P.; Jafari, R. Enabling Effective Programming and Flexible Management of Efficient Body Sensor Network Applications. *IEEE Trans. Hum. Mach. Syst.* **2013**, *43*, 115–133. [CrossRef]
3. Sodhro, A.H.; Luo, Z.; Sangaiah, A.K.; Baik, S.W. Mobile edge computing based QoS optimization in medical healthcare applications. *Int. J. Inf. Manag.* **2018**. [CrossRef]
4. Sodhro, A.H.; Pirbhulal, S.; Qaraqe, M.; Lohano, S.; Sodhro, G.H.; Junejo, N.U.; Luo, Z. Power control algorithms for media transmission in remote healthcare systems. *IEEE Access* **2018**, *6*, 42384–42393. [CrossRef]
5. Cao, J.; Li, W.; Ma, C.; Tao, Z. Optimizing multi-sensor deployment via ensemble pruning for wearable activity recognition. *Inf. Fusion* **2018**, *41*, 68–79. [CrossRef]
6. Galzarano, S.; Giannantonio, R.; Liotta, A.; Fortino, G. A task-oriented framework for networked wearable computing. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 621–638. [CrossRef]
7. Abebe, G.; Cavallaro, A. Inertial-Vision: Cross-Domain Knowledge Transfer for Wearable Sensors. In Proceedings of the 2017 International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1392–1400.
8. Fortino, G.; Galzarano, S.; Gravina, R.; Li, W. A framework for collaborative computing and multi-sensor data fusion in body sensor networks. *Inf. Fusion* **2015**, *22*, 50–70. [CrossRef]
9. Yang, X.; Tian, Y. Super Normal Vector for Human Activity Recognition with Depth Cameras. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1028–1039. [CrossRef]
10. Bux, A.; Angelov, P.; Habib, Z. Vision Based Human Activity Recognition: A Review. *Adv. Comput. Intell. Syst.* **2017**, *513*, 341–371.
11. Zhao, H.; Ding, Y.; Yu, B.; Jiang, C.; Zhang, W. Design and implementation of Peking Opera action scoring system based on human skeleton information. *MATEC Web Conf.* **2018**, *232*, 01026. [CrossRef]
12. Vicon. Available online: <https://www.vicon.com/> (accessed on 15 January 2019).
13. Optotrak. Available online: <https://www.ndigital.com/> (accessed on 15 January 2019).
14. Hong, J.H.; Ramos, J.; Dey, A.K. Toward Personalized Activity Recognition Systems With a Semipopulation Approach. *IEEE Trans. Hum. Mach. Syst.* **2016**, *46*, 101–112. [CrossRef]
15. Xsens. Available online: <https://www.xsens.com/> (accessed on 15 January 2019).
16. Invensense. Available online: <https://www.invensense.com/> (accessed on 15 January 2019).
17. VMSens. Available online: <https://www.vmsens.com> (accessed on 15 January 2019).
18. Noitom. Available online: <http://www.noitom.com.cn/> (accessed on 15 January 2019).
19. Decawave. Available online: <https://www.decawave.com/> (accessed on 15 January 2019).
20. Cornacchia, M.; Ozcan, K.; Zheng, Y.; Velipasalar, S. A Survey on Activity Detection and Classification Using Wearable Sensors. *IEEE Sens. J.* **2017**, *17*, 386–403. [CrossRef]
21. Bao, S.D.; Meng, X.L.; Xiao, W.; Zhang, Z.Q. Fusion of Inertial/Magnetic Sensor Measurements and Map Information for Pedestrian Tracking. *Sensors* **2017**, *17*, 340. [CrossRef]
22. Xu, C.; He, J.; Zhang, X.; Yao, C.; Tseng, P.H. Geometrical Kinematic Modeling on Human Motion using Method of Multi-Sensor Fusion. *Inf. Fusion* **2017**, *41*, 243–254. [CrossRef]

23. Zihajehzadeh, S.; Park, E.J. A Novel Biomechanical Model-Aided IMU/UWB Fusion for Magnetometer-Free Lower Body Motion Capture. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 927–938. [CrossRef]
24. Qiu, S.; Wang, Z.; Zhao, H.; Hu, H. Using Distributed Wearable Sensors to Measure and Evaluate Human Lower Limb Motions. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 939–950. [CrossRef]
25. Fullerton, E.; Heller, B.; Munoz-Organero, M. Recognising human activity in free-living using multiple body-worn accelerometers. *IEEE Sens. J.* **2017**, *17*, 5290–5297. [CrossRef]
26. Chen, C.; Jafari, R.; Kehtarnavaz, N. A survey of depth and inertial sensor fusion for human action recognition. *Multimed. Tools Appl.* **2017**, *76*, 4405–4425. [CrossRef]
27. Nilsson, J.O.; Zachariah, D.; Skog, I.; Händel, P. Cooperative localization by dual foot-mounted inertial sensors and inter-agent ranging. *Eur. J. Adv. Signal Process.* **2013**, *2013*, 164. [CrossRef]
28. Sodhro, A.H.; Shaikh, F.K.; Pirbhulal, S.; Lodro, M.M.; Shah, M.A. Medical-QoS based telemedicine service selection using analytic hierarchy process. In *Handbook of Large-Scale Distributed Computing in Smart Healthcare*; Springer: Cham, Switzerland, 2017; pp. 589–609.
29. Lodro, M.M.; Majeed, N.; Khuwaja, A.A.; Sodhro, A.H.; Greedy, S. Statistical channel modelling of 5G mmWave MIMO wireless communication. In Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 3–4 March 2018; pp. 1–5.
30. Magsi, H.; Sodhro, A.H.; Chachar, F.A.; Abro, S.A.K.; Sodhro, G.H.; Pirbhulal, S. Evolution of 5G in Internet of medical things. In Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 3–4 March 2018.
31. Sodhro, A.H.; Pirbhulal, S.; Sodhro, G.H.; Gurtov, A.; Muzammal, M.; Luo, Z. A joint transmission power control and duty-cycle approach for smart healthcare system. *IEEE Sens. J.* **2018**. [CrossRef]
32. Kinect. Available online: <https://developer.microsoft.com/en-us/windows/kinect> (accessed on 15 January 2019).
33. Iyengar, S.; Bonda, F.T.; Gravina, R.; Guerrieri, A.; Fortino, G.; Sangiovanni-Vincentelli, A. A framework for creating healthcare monitoring applications using wireless body sensor networks. In Proceedings of the ICST 3rd International Conference on Body Area Networks, Tempe, AZ, USA, 13–17 March 2008.
34. Anwary, A.R.; Yu, H.; Vassallo, M. Optimal Foot Location for Placing Wearable IMU Sensors and Automatic Feature Extraction for Gait Analysis. *IEEE Sens. J.* **2018**, *18*, 2555–2567. [CrossRef]
35. Ghasemzadeh, H.; Jafari, R. Physical Movement Monitoring Using Body Sensor Networks: A Phonological Approach to Construct Spatial Decision Trees. *IEEE Trans. Ind. Inf.* **2011**, *7*, 66–77. [CrossRef]
36. Yun, X.; Bachmann, E.R. Design, Implementation, and Experimental Results of a Quaternion-Based Kalman Filter for Human Body Motion Tracking. *IEEE Int. Conf. Robot. Autom.* **2006**, *22*, 1216–1227. [CrossRef]
37. Zhao, S.; Shmaliy, Y.S.; Shi, P.; Ahn, C.K. Fusion Kalman/UFIR filter for state estimation with uncertain parameters and noise statistics. *IEEE Trans. Ind. Electr.* **2017**, *64*, 3075–3083. [CrossRef]
38. Briese, D.; Kunze, H.; Rose, G. UWB localization using adaptive covariance Kalman Filter based on sensor fusion. In Proceedings of the 2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB), Salamanca, Spain, 12–15 September 2017.
39. Kim, H.; Liu, B.; Goh, C.Y.; Lee, S.; Myung, H. Robust vehicle localization using entropy-weighted particle filter-based data fusion of vertical and road intensity information for a large scale urban area. *IEEE Robot. Autom. Lett.* **2017**, *2*, 1518–1524. [CrossRef]
40. Bai, F.; Vidal-Calleja, T.; Huang, S. Robust Incremental SLAM Under Constrained Optimization Formulation. *IEEE Robot. Autom. Lett.* **2018**, *3*, 1207–1214. [CrossRef]
41. Sebastián, P.S.J.; Virtanen, T.; Garcia-Molla, V.M.; Vidal, A.M. Analysis of an efficient parallel implementation of active-set Newton algorithm. *J. Supercomput.* **2018**. [CrossRef]
42. Liu, Z.; Zhang, L.; Liu, Q.; Yin, Y.; Cheng, L.; Zimmermann, R. Fusion of magnetic and visual sensors for indoor localization: Infrastructure-free and more effective. *IEEE Trans. Multimed.* **2017**, *19*, 874–888. [CrossRef]
43. Tichavsky, P.; Muravchik, C.H.; Nehorai, A. Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Process.* **1998**, *46*, 1386–1396. [CrossRef]
44. Qi, Y.; Kobayashi, H.; Suda, H. Analysis of wireless geolocation in a non-line-of-sight environment. *IEEE Trans. Wirel. Commun.* **2006**, *5*, 672–681.

45. Yang, Z.; Liu, Y. Quality of Trilateration: Confidence Based Iterative Localization. *IEEE Trans. Parallel Distrib. Syst.* **2010**, *21*, 631–640. [[CrossRef](#)]
46. Geng, Y.; Pahlavan, K. Design, implementation, and fundamental limits of image and RF based wireless capsule endoscopy hybrid localization. *IEEE Trans. Mob. Comput.* **2016**, *15*, 1951–1964. [[CrossRef](#)]
47. Zhang, H.; Dufour, F.; Anselmi, J.; Laneuville, D.; Nègre, A. Piecewise optimal trajectories of observer for bearings-only tracking by quantization. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017; pp. 1–7.
48. Fisher/Information. Available online: https://en.wikipedia.org/wiki/Fisher_information (accessed on 15 January 2019).
49. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 2012.
50. Gasparini, M. Markov Chain Monte Carlo in Practice. *Technometrics* **1999**, *39*, 338. [[CrossRef](#)]
51. Nilsson, J.O.; Gupta, A.K.; Handel, P. Foot-mounted inertial navigation made easy. In Proceedings of the 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, Korea, 27–30 October 2014.
52. Fan, Q.; Zhang, H.; Sun, Y.; Zhu, Y.; Zhuang, X.; Jia, J.; Zhang, P. An Optimal Enhanced Kalman Filter for a ZUPT-Aided Pedestrian Positioning Coupling Model. *Sensors* **2018**, *18*, 1404. [[CrossRef](#)] [[PubMed](#)]
53. Xu, C.; He, J.; Zhang, X.; Tseng, P.H.; Duan, S. Toward Near-Ground Localization: Modeling and Applications for TOA Ranging Error. *IEEE Trans. Antennas Propag.* **2017**, *65*, 5658–5662. [[CrossRef](#)]
54. Deng, C.Y. A generalization of the ShermanMorrisonWoodbury formula. *Appl. Math. Lett.* **2011**, *24*, 1561–1564. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).