



HAL
open science

Sparse Visual-Inertial Measurement Units Placement for Gait Kinematics Assessment

Randa Mallat, Vincent Bonnet, Raphaël Dumas, Mohamed Adjel, Gentiane Venture, Mohamad Khalil, Samer Mohammed

► **To cite this version:**

Randa Mallat, Vincent Bonnet, Raphaël Dumas, Mohamed Adjel, Gentiane Venture, et al.. Sparse Visual-Inertial Measurement Units Placement for Gait Kinematics Assessment. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29, pp.1300-1311. 10.1109/TNSRE.2021.3089873 . hal-03313483

HAL Id: hal-03313483

<https://hal.science/hal-03313483v1>

Submitted on 4 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sparse Visual-Inertial Measurement Units Placement for Gait Kinematics Assessment

Randa Mallat, Vincent Bonnet, Raphael Dumas, Mohamed Adjel, Gentiane Venture, Mohamad Khalil, Samer Mohammed

Abstract—This study investigates the possibility of estimating lower-limb joint kinematics and meaningful performance indexes for physiotherapists, during gait on a treadmill based on data collected from a sparse placement of new Visual Inertial Measurement Units (VIMU) and the use of an Extended Kalman Filter (EKF). The proposed EKF takes advantage of the biomechanics of the human body and of the investigated task to reduce sensor inaccuracies. Two state-vector formulations, one based on the use of constant acceleration model and one based on Fourier series, and the tuning of their corresponding parameters were analyzed. The constant acceleration model, due to its inherent inconsistency for human motion, required a cumbersome optimisation process and needed the a-priori knowledge of reference joint trajectories for EKF parameters tuning. On the other hand, the Fourier series formulation could be used without a specific parameters tuning process. In both cases, the average root mean square difference and correlation coefficient between the estimated joint angles and those reconstructed with a reference stereophotogrammetric system was 3.5deg and 0.70, respectively. Moreover, the stride lengths were estimated with a normalized root mean square difference inferior to 2% when using the forward kinematics model receiving as input the estimated joint angles. The popular gait deviation index was also estimated and showed similar results very close to 100, using both the proposed method and the reference stereophotogrammetric system. Such consistency was obtained using only three wireless and affordable VIMU located at the pelvis and both heels and tracked using two affordable RGB cameras. Being further easy-to-use and suitable for applications taking place outside of the laboratory, the proposed method thus represents a good compromise between accurate reference stereophotogrammetric systems and markerless ones for which accuracy is still under debate.

Index Terms—Extended Kalman Filter, Gait rehabilitation, Inertial Measurement Unit, Augmented Reality Markers.

I. INTRODUCTION

GAIT analysis is the most popular method used to quantify movement disorders and to assess rehabilitation effects over time [1]. It is a clinical routine used in broad variety of pathologies in rheumatology, orthopedics, endocrinology, and neurology [2]. When assessing gait, kinematics parameters such as joint angles, that are the basis of most analyses, or spatiotemporal parameters (e.g., velocity and stride length)

R. Mallat, V. Bonnet, M. Adjel, and S. Mohammed are with the Univ Paris Est Creteil, LISSI, F-94400 Vitry, France.

E-mail: vincent.bonnet@u-pec.fr
V. Bonnet is also with the LAAS - CNRS, CNRS, UPR 8001, Toulouse, France.

R. Dumas is with Université de Lyon, Université Gustave Eiffel, LBMC UMR T9406, F69622, Lyon, France.

G. Venture is with Tokyo University of Agriculture and Technology, Tokyo, Japan.

R. Mallat and M. Khalil are with the Faculty of Engineering and Azm Center at the Lebanese University.

are often used. When accurately estimated, those parameters can be combined to build assessment indicators to assist the therapists in their diagnosis. For example, the Gait Deviation Index (GDI) introduced by Schwartz et al. is increasingly used by the clinical community to estimate the level of gait pathology among disabled persons [3]. The GDI compares gait kinematics data of a given patient to a healthy group, by estimating the distance between healthy strides and patient's strides. This distance (with respect to the healthy group), which reflects the pathology level, requires an estimation of the kinematics parameters to be assessed properly. Stereophotogrammetric Systems (SS) are considered as the reference tools in kinematics motion analysis. However, their use require a large financial investment and a complex experimental protocol. An alternative solution proposed by Cappozzo was called the minimum measured input model approach, that is aiming at maximizing the functional information extracted from simplified and affordable experimental protocols [4].

A. Related works

In this context, mainly two technologies were proposed: Inertial Measurement Unit (IMU) and RGB/RGB-Depth camera-based systems. Small, low-cost IMUs measure three accelerations and three angular velocities. Thus, in principle, it is possible to estimate their position and orientation through double and single integration, respectively. Unfortunately, drift jeopardizes the time integration of raw signals [5]. To overcome this problem, Kalman adaptive filters are often proposed in the literature to ensure real-time estimates of the IMU orientation by fusing accelerometer and gyroscope data. Since only two orientation angles can be obtained with this approach [6], several studies have proposed to use at least one IMU for each segment of interest together with a multi-body model representing the human skeleton as a mean to reduce drift-influence. For example, El Gohary et al. proposed to use a zero velocity update approach with an Unscented Kalman Filter relying on a kinematics model for arm motions tracking [7]. Remarkable results were obtained when operating with a rigid robotic arm so that human soft-tissue-artifact and sensor-to-segment calibration errors were not considered in this study. To reduce the sensors count, Sy et al. proposed recently a method to estimate lower limbs 3D joint angles during gait [8]. Despite a very interesting approach based on the use of an Extended Kalman Filter (EKF) including assumptions and constraints such as rigid body constraints, zero velocity update at the foot contact, assuming flat floor or making use of the total pelvis height, the presented results displayed a large Root Mean

Square Difference (RMSD) superior to 10deg when compared to the same estimates using SS. These relatively poor results might be due to the fact that a proper sensor-to-segment calibration was missing and that a constant acceleration model was used. Indeed, IMUs require a calibration phase to align each sensor with respect to its segment. Static and functional calibration techniques based on a set of pre-defined postures and/or motions are commonly used in literature. However, these methods are time-consuming, might be limited by the subject's ability to perform pre-defined postures or motions, and thus can be hardly used with some population or in clinical settings for example [9]. Large errors might also be due to sub-optimal weights tuning leading to an emphasized influence of IMU drift on joint angle estimates. Joukov et al. showed the influence of parameter's tuning when using an adaptive filter such as the EKF [10]. Moreover, during gait, they showed the superiority of using Fourier series over using standard constant acceleration model of joint temporal evolution. The assumption that the joint trajectories can be modelled by using low-order Fourier series has already been made in the literature [10]–[12]. Fourier representation was also largely used for spatio-temporal parameters assessment using IMUs [13]. Fourier series coefficients were identified using a feedback adaptive frequency phase oscillator. As Joukov et al. discussed, the learning rate and parameters tuning of the oscillator was decisive in the performance of their method [10]. Nevertheless, by performing a grid search, it was possible to fine-tune the filter and obtain a low RMSD of 2.4deg for the hip and knee joint angles. However, it should be noted that the proposed method could not estimate ankle angles, required a cumbersome sensor-to-segment calibration process and was based on one IMU per investigated segment. To cope with the IMU drawbacks, some authors have proposed to use camera based systems. Recently, markerless visual motion capture based on a RGB camera and machine learning algorithms were developed for estimating human motion in real-life scenarios [14], [15]. However, such approaches are not yet real-time or accurate enough to reliably estimate dependent subject's joint angles and thus are not ready to be used yet in clinical settings for example. Moreover such methods need often to be trained on large scale datasets requiring a large computational power. Devices, based on RGB-Depth sensors such as the Kinect camera (Microsoft), with embedded skeleton tracking algorithms are available. However, it was shown that the Kinect camera is not accurate enough for rehabilitation assessment mainly due to segment length variations [16]. For gait analysis, the Kinect sensor can only estimate timing and spatial characteristics [17] as joint angles are not reliable enough for a consistent analysis. Numerous methods have been developed in the literature to improve kinematic estimates of RGB-D like sensors. Most of them add the use of a multi-body model to obtain a more robust joint angle estimate [15], [16]. However, acceptable differences can only be obtained for motions of very large amplitude [16].

Another camera based approach is homography, which is used in Augmented Reality (AR) to accurately estimate 3D rotations and translations of a given AR marker after a proper,

yet simple, camera calibration process.

Nagy et al. proposed the use of AR markers for gait analysis [18]. They have used a commercial Multi-body Kinematics Optimization (MKO) approach that was fed with segments 3D pose obtained from very large AR markers that were located at each investigated segments. They displayed a RMSD of 2.3cm for the step length, and a joint angle estimate varying from 2.5deg to 6.7deg. The authors state that they have chosen to use such large AR markers to avoid occlusion and undetected markers. In a recent survey [19], it was suggested that merging RGB-Depth and IMU data could prevent occlusions and improve joint angle estimate accuracy [20]. Ahmed et al. for example were able to estimate head and feet trajectories during overground walking, with a RMSD that was less than 2% of the walked distance, based on the fusion of only three sensors: a head mounted IMU-camera and two feet mounted IMU [21]. However, the joint kinematics were not assessed in this study. Moreover, it is worthy noting that MKO and EKF were alternatively used with IMU [10] and AR markers [18], [22] (as well as with SS markers [23]). Both methods seem to provide similar results however the MKO can hardly be implemented in real-time [24].

In this context, it seems that there is a consensus in the community on the fact that a multi-body model based approach should be used to perform inverse kinematics with a sparse placement of sensors. Fusing camera based information, obtained from AR markers, and IMU data could lead to reliable and robust 3D sensor pose estimate. Proper sensor-to-segment calibration and EKF parameters tuning are also essential to provide accurate joint kinematics estimation. Our group has already proposed such an approach in a recent study [22] but it was dealing with simple arm tasks. Moreover, it required one sensor module per investigated segment. Nevertheless it allowed to show that it is possible to reduce IMU drift influence and to cope with visual occlusion problems.

B. Contribution

The objective of the present study is to demonstrate the validity of using sparse calibrated newly designed Visual-Inertial Measurement Units (VIMU) and adaptive filtering, during treadmill gait, to estimate relevant kinematics quantities that may be of great benefit to the clinical community. Moreover, besides new constraints implemented in the EKF, two formulations of the proposed state-vector are studied and compared: a classical constant acceleration model associated to an optimal tuning process of the EKF's parameters, and a model based on Fourier series that allows a simple tuning of the EKF's parameters based on a priori knowledge. Thus one of the sub-objective of this study is to assess if by having a more realistic state variable evolution model the tuning of the EKF parameters will be simplified.

II. METHOD

The overall principle of the proposed approach is depicted in Fig.1, and is based on the simultaneous use of affordable IMUs, incorporating only 3D accelerometers and gyroscopes sensors, i.e., no magnetometers, and a set of AR markers

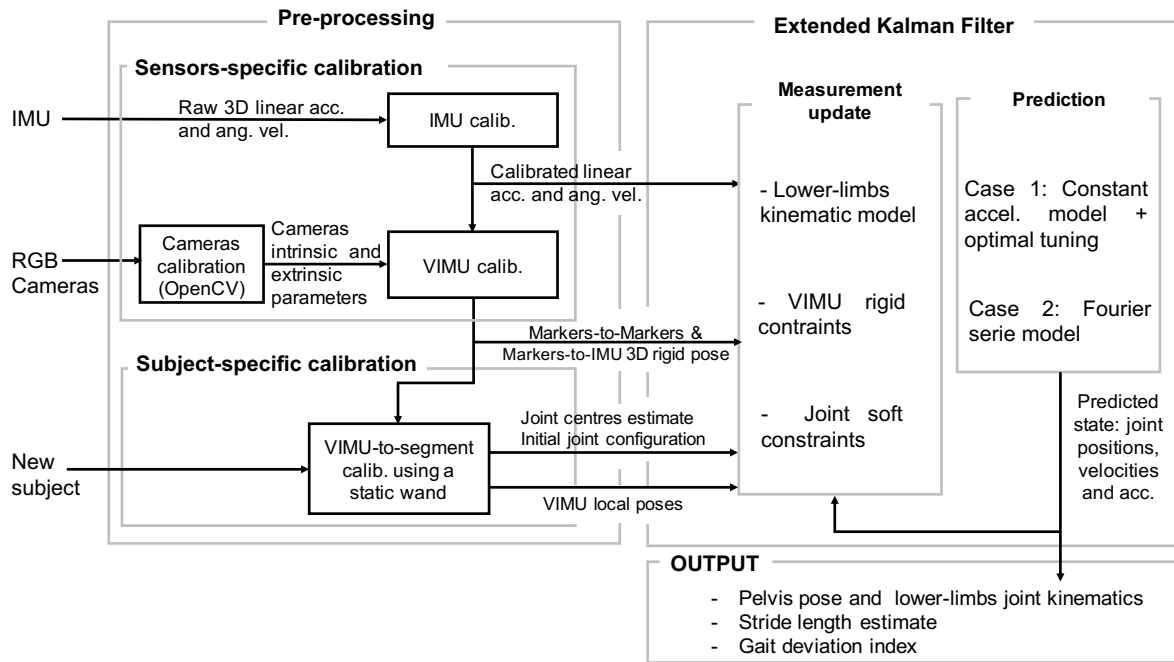


Fig. 1. Overview of the proposed approach based on Visual-Inertial Measurement Units (VIMU) and an Extended Kalman Filter (EKF) taking advantage of the human body and task biomechanics.

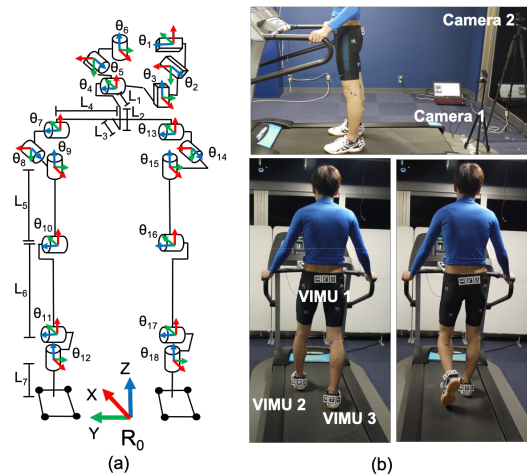


Fig. 2. (a) Lower-limbs multi-body model composed of N_S segments, of lengths $L_i, i = 1, \dots, N_S$, and N_J Degrees-of-Freedoms (DoF) denoted $\theta_i, i = 1, \dots, N_J$. (b) Experimental setup, three VIMU attached to the pelvis, left, and right heels in a bracelet form are collecting data during gait thanks to two RGB cameras located behind the treadmill.

tracked with two affordable RGB cameras. The raw data collected from both IMUs and AR markers require a prior and unique calibration procedure that is based on a process detailed in our previous study [22] and briefly described in section II.B. A set of three AR markers is mounted on the top of each IMU sensor to form a so-called VIMU (see Fig. 3). Then, a practical subject-specific sensor-to-segment static calibration (section II.B.2) is done to define each VIMU pose in the segment coordinate systems (subject's segment axes and joint centres are therefore defined at this step). Finally, an EKF based on an eighteen Degrees-of-Freedom (DoF) multi-body kinematics

model that is described in section II.A, fuses all measured quantities to estimate lower-limbs joint angles. The tuning of the EKF's parameters along with its state-vector expressed in two different versions with respect to the evolution of the joint trajectories are also compared and discussed (section II.C and section II.D).

A. Lower-limbs Multi-Body Model

A multi-body model of the lower-limbs is devised to relate the joint kinematics with the three VIMU measurements. It consists of $N_S = 7$ segments articulated with $N_J = 18$ DoF. As represented in Fig. 2.a, it is composed of two spherical joints (hip), two hinge joints (knee) and two universal joints (ankle). Moreover, the model floating base is located with respect to a global coordinate system (R_0) using three prismatic and three rotational virtual joints. Three VIMUs are assumed to be rigidly attached to the sacrum, and to the left and right heels, respectively (Fig 2.b). A VIMU is composed of three AR markers and a single IMU (see Fig.3 and Fig.4). The pose (position and orientation) of each VIMU marker is directly estimated by the off-the-shelf ArUco library [46]. The orientation provided by ArUco is expressed using the quaternion formalism as presented by Diebel [25]. Thus, each VIMU allows measuring the absolute 3D position \mathbf{p}_v^0 and orientation (quaternion) \mathbf{q}_v^0 with respect to R_0 of each AR marker as well as the 3D linear acceleration \mathbf{a}_v^v and angular velocity $\boldsymbol{\Omega}_v^v$ with respect to the IMU coordinate system. Three coordinate systems are therefore used to represent VIMU different elements: a coordinate system corresponding to both IMU and central AR marker and two additional coordinate systems for the left and right AR markers, respectively. The forward kinematics model (FKM) and its first and second

derivative (\mathbf{J} and $\dot{\mathbf{J}}$) relating joint kinematics to the VIMU measurements are given as follows:

$$\begin{aligned} \mathbf{p}_v^0 &= FKM(\theta) \\ \mathbf{q}_v^0 &= f(\mathbf{R}_v^0) \\ \boldsymbol{\Omega}_v^0 &= (\mathbf{R}_v^0)^T \mathbf{J}_R(\theta)\dot{\theta} \\ \mathbf{a}_v^0 &= (\mathbf{R}_v^0)^T (\mathbf{J}_P(\theta)\ddot{\theta} + \dot{\mathbf{J}}_P(\theta)\dot{\theta}) + \mathbf{b}_a \end{aligned} \quad (1)$$

where θ , $\dot{\theta}$, $\ddot{\theta}$ are the $(N_J \times 1)$ joints positions, velocities and accelerations vectors, respectively. \mathbf{R}_v^0 (3×3) is the rotation matrix expressing the VIMU's central AR marker orientation in R_0 . It is converted into a (4×1) quaternion vector \mathbf{q}_v^0 through a custom mapping function $f(\cdot)$. Details on the implementation of $f(\cdot)$ can be found in section 6.5 of the review paper proposed by Diebel [25]. The Jacobian matrix (6×3) is composed of $\mathbf{J} = [\mathbf{J}_P \quad \mathbf{J}_R]$, and \mathbf{b}_a is a (3×1) acceleration bias vector.

B. Calibration

Two easy-to-use calibration procedures are required to use the raw measurements effectively:

- a calibration for each VIMU, independent of the subject,
- a subject-specific sensor-to-segment calibration.

On the one hand, the VIMU calibration is done without participation of the subject by positioning and moving the unit in the field of view of the RGB camera. On the other hand, thanks to the use of AR markers, an anatomical calibration of the VIMU can simply be done [26].

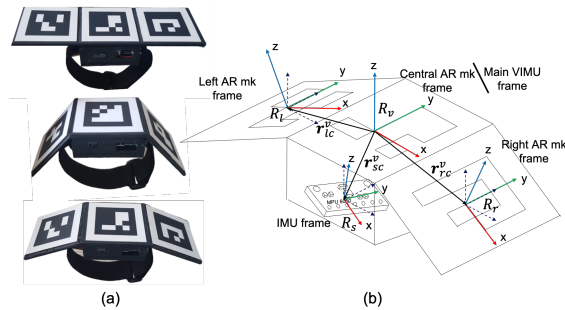


Fig. 3. (a) The VIMU prototype with three different bracelet arrangements to fit different subjects and/or segment sections. (b) Description of the coordinate systems associated to a VIMU. The coordinate system of the IMU sensor, left, and right AR markers are all linked to the central marker coordinate system through rotation matrices \mathbf{R} and a translation vectors \mathbf{r} .

1) *VIMU calibration*: The gyroscope and accelerometer were calibrated, without the use of any external device, as proposed by Tedaldi et al. to take into account inaccuracies in scaling factors, misalignment and non-orthogonality of axes, and non-zero biases [27]. Moreover, as shown in our previous study [22], the coordinate systems of the IMU and of the central AR markers were aligned thanks to a dynamic calibration step making use of centripetal acceleration, to calculate the rotation matrix \mathbf{R}_v^0 (3×3). Two rigid transformations also exist between the side coordinate systems of the AR markers and of the central one:

$$\begin{aligned} \mathbf{R}_{ic}^v &= (\mathbf{R}_v^0)^T \mathbf{R}_i^0 \\ \mathbf{q}_{ic}^v &= f(\mathbf{R}_{ic}^v) \\ \mathbf{r}_{ic}^v &= (\mathbf{R}_v^0)^T (\mathbf{r}_i^0 - \mathbf{r}_v^0) \end{aligned} \quad (2)$$

where \mathbf{R}_{ic}^v (3×3), \mathbf{q}_{ic}^v (4×1), \mathbf{r}_{ic}^v (3×1) represent the rotation matrix, its corresponding quaternion formulation, and position vector of $i=l$ (left) or r (right) AR marker expressed in the central one and obtained after calibration, respectively. These relative transformations were easily determined using 50 different static postures of the VIMU and solving an over-determined system [28]. These rigid transformations between the coordinate systems of the AR markers will be considered as rigid body constraints in the EKF (section II.C). Finally, a VIMU is fully calibrated given that all its measurements are expressed with respect to the coordinate system attached to the central AR marker frame R_v .

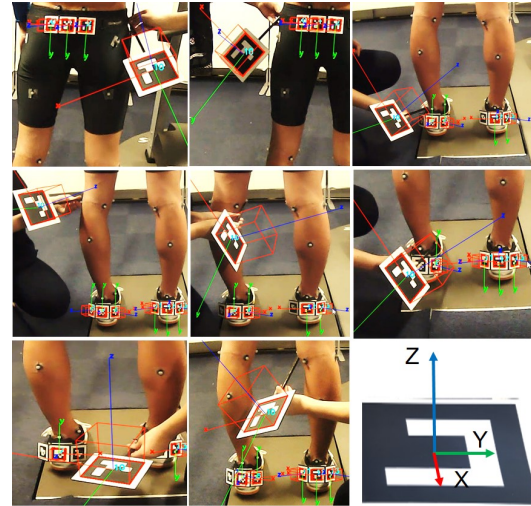


Fig. 4. Example of lower-limbs anatomical landmarks pin-pointed using a calibration wand based on an AR marker.

2) *Sensor-to-Segment Calibration*: As exemplify in Fig.4 the information from an additional AR marker mounted on a large calibration wand, making its detection accurate and robust, can be used to pin-point anatomical landmarks similarly to what is done when using a SS [26]. The calibration wand is used to pin-point eleven lower-limbs anatomical landmarks in R_0 . These are the midpoint between the postero-superior illiac spines, the right and left antero-superior illiac spines, the lateral and medial femoral epicondyles, and the lateral and medial malleolus. This operation can be done in less than a minute. Joint center positions are then calculated using classical linear regression methods [29]. The norm between consecutive joint center positions is used to determine the segments lengths in the multi-body model (section II.A, Fig. 2). Note that this sensor-to-segment calibration is valid as long as the VIMUs location on their corresponding segments are not modified. Otherwise, it has to be repeated.

C. Extended Kalman Filter

An EKF is proposed to estimate lower-limbs joint kinematics, gathered in the state-vector \mathbf{x} , that minimizes at each time step k the least-square difference between the VIMU measurements, gathered in \mathbf{y} , and their estimate from the measurement model h , gathered in \mathbf{h} , including the estimate of the VIMUs 3D positions \mathbf{p}_v^0 , orientations \mathbf{q}_v^0 , linear accelerations \mathbf{a}_v^0 and angular velocities $\boldsymbol{\Omega}_v^0$ (section II.A).

$$\begin{aligned} \mathbf{h}_k &= [\mathbf{p}_{vk}^0 \quad \mathbf{q}_{vk}^0 \quad \mathbf{a}_{vk}^v \quad \boldsymbol{\Omega}_{vk}^v \quad \mathbf{h}_{c1k} \quad \mathbf{h}_{c2k}]^T \\ \mathbf{h}_{c1k} &= [(\theta_{ik} - \theta_M)^2]^T \quad \text{with } i = 7, \dots, N_J \\ \mathbf{h}_{c2k} &= [(\mathbf{r}_l^v - \mathbf{r}_{lc}^v)^2 \quad (\mathbf{q}_l^v - \mathbf{q}_{lc}^v)^2 \quad (\mathbf{r}_r^v - \mathbf{r}_{rc}^v)^2 \quad (\mathbf{q}_r^v - \mathbf{q}_{rc}^v)^2]^T \end{aligned} \quad (3)$$

The vector of estimated VIMUs measurements \mathbf{h} is based on the forward kinematics of the multi-body model but also incorporates two type of constraints terms (c1 and c2). The elements of \mathbf{h}_{c1} ($N_J \times 1$) are used to solve kinematics indetermination. Since the number of VIMU is reduced, i.e. a single VIMU per kinematics chain, there exist multiple kinematics solutions leading to the same VIMU measurements in the Cartesian space. However, there is not an infinite number of solutions leading to the same VIMU pose as the proposed leg model has 6DoF similarly to the number of estimated quantities from a VIMU. This kinematics indetermination is well known as the elbow-up/elbow-down problem in robotics. It has been solved in different ways. In this study, we propose to push the joint solution toward the average joint angle value, θ_M , reported for walking in the literature [30], [31]: 13deg, 8deg, 11deg, 16deg, 1deg, 6deg for $\theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}$ angles of the left leg and, symmetrically for the right leg, respectively. In the context of the EKF, this is done by adding $N_J - 6$ virtual measurements aiming to minimise the square difference with θ_M at each sample of time by setting the corresponding measurements in \mathbf{y} to zero. This is equivalent to minimizing a cost function in an optimization process [32].

The elements in \mathbf{h}_{c2} (42×1) additionally ensure rigid body constraints between the three AR markers in the VIMU. The position and orientation of the right and left AR markers with respect to the central AR marker should match their values obtained during the VIMU calibration (section II.B).

Once the measurement model was established, the joint kinematics estimation was conducted, first, through the prediction of the a priori state-vector $\hat{\mathbf{x}}^-$ thanks to a state evolution model f , that is approximated by a linear form \mathbf{F} , and a process covariance matrix \mathbf{Q} :

$$\begin{aligned} \hat{\mathbf{x}}_k^- &= \mathbf{F}\hat{\mathbf{x}}_{k-1} \\ \mathbf{P}_k^- &= \mathbf{F}\mathbf{P}_{k-1}^-\mathbf{F}^T + \mathbf{Q} \end{aligned} \quad (4)$$

where \mathbf{P}_k^- is the a priori estimation of the error covariance matrix.

Two cases were investigated for the state-vector formulation:

- *case 1 EKF based on a Constant Acceleration (EKF-CA) model*: A first classical state-vector formulation proposes to estimate the joint angles, velocities and accelerations and assumed a constant acceleration model [23]:

$$\begin{aligned} \mathbf{x}_{ik} &= [\theta_{ik} \quad \dot{\theta}_{ik} \quad \ddot{\theta}_{ik} \quad \mathbf{b}_{ak}]^T \quad \text{with } i = 1, \dots, N_J \\ \mathbf{F} &= \begin{bmatrix} \mathbf{I}_{N_J \times N_J} & \Delta T_s \mathbf{I}_{N_J \times N_J} & \frac{\Delta T_s^2}{2} \mathbf{I}_{N_J \times N_J} & \mathbf{0}_{N_J \times 9} \\ \mathbf{0}_{N_J \times N_J} & \mathbf{I}_{N_J \times N_J} & \Delta T_s \mathbf{I}_{N_J \times N_J} & \mathbf{0}_{N_J \times 9} \\ \mathbf{0}_{N_J \times N_J} & \mathbf{0}_{N_J \times N_J} & \mathbf{I}_{N_J \times N_J} & \mathbf{0}_{N_J \times 9} \\ \mathbf{0}_{9 \times 3 N_J} & \mathbf{0}_{9 \times 3 N_J} & \mathbf{0}_{9 \times 3 N_J} & \mathbf{I}_{9 \times 9} \end{bmatrix} \end{aligned} \quad (5)$$

where T_s , $\mathbf{0}$ and \mathbf{I} are the sampling time, and the null and identity matrices, respectively. To reduce the drift effect, a

random acceleration bias \mathbf{b}_a (9×1) was added to better track the 3D measured accelerations of each VIMU (see Eq. 1) [7].

- *case 2 EKF based on a Fourier Series (EKF-FS) model*: Since gait can be assumed to be a quasi-periodic task, the angle of the i^{th} joint can be represented using a Fourier series:

$$\theta_{ik} = \mathbf{a}_{i0} + \sum_{n=1}^{N_H} (\mathbf{a}_{in} \cos(n\omega k T_s) + \mathbf{b}_{in} \sin(n\omega k T_s)) \quad (6)$$

where N_H is the number of harmonics, and ω is the main motion pulsation.

As described in the experimental validation (see section III), the gait activity takes place on a treadmill at constant velocity. Consequently, Fourier series coefficients are expected to converge towards near constant values and identity matrix can be used as state-vector evolution model. The state-vector and its evolution model were defined as follows:

$$\begin{aligned} \mathbf{x}_{ik} &= [\mathbf{a}_{i0} \quad \mathbf{a}_{in} \quad \mathbf{b}_{in} \quad \omega_k \quad \mathbf{b}_{ak}]^T \\ \text{with } i &= 1, \dots, N_J \quad \text{and } n = 1, \dots, N_H \end{aligned} \quad (7)$$

$$\mathbf{x}_{k+1} = \mathbf{I}_{N_x \times N_x} \mathbf{x}_k$$

with $N_x = (1 + 2N_H)N_J + 1 + 9$.

Secondly, depending on the measurement error, \mathbf{v} , and the measurement covariance matrix, \mathbf{R} , the Kalman gain \mathbf{K} was calculated to update the state-vector:

$$\begin{aligned} \mathbf{v}_k &= \mathbf{y}_k - h(\hat{\mathbf{x}}_k^-) \\ \mathbf{S}_k &= \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \\ \mathbf{K}_k &= \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{S}_k^{-1} \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k \mathbf{v}_k \\ \mathbf{P}_k &= \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T + (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T \end{aligned} \quad (8)$$

where \mathbf{H} is the Jacobian matrix $\frac{\partial h}{\partial \mathbf{x}}$ and \mathbf{S} is the measurement covariance matrix.

D. EKF Parameters Tuning

The proposed EKF requires the tuning of the following elements: the initial state-vector value $\hat{\mathbf{x}}_0$, the initial estimation of the error covariance matrix \mathbf{P}_0 , the process noise covariance matrix \mathbf{Q} and the measurement noise covariance matrix \mathbf{R} .

For the initial state-vector and error covariance matrices, it has been shown that their values mainly affect the initial part of the estimation [33]. Thus, $\hat{\mathbf{x}}_0$ was set with the result of the inverse geometric model calculated from the AR markers data collected at the first sample of time. \mathbf{P}_0 was set equal to the identity matrix giving the same influence to all joints.

The tuning of the elements of \mathbf{Q} and \mathbf{R} is more sensitive as they impact the filter convergence rate and stability. The elements of \mathbf{R} corresponding to the IMU data can be experimentally determined. However, the accuracy of the visual measurements is variable as a function of different aspects such as marker occlusion, motion blur, or distance and orientation with respect to the camera and thus is difficult to be assessed. The elements of \mathbf{R} corresponding to the constraints

c1 and c2 (section II.C) must also be tuned leading to a large number of parameters. Moreover, the ratio between the tuning of the elements of \mathbf{R} and \mathbf{Q} , that is more sensitive and task and joint dependent [23], affects the filter performance [33]. The tuning of \mathbf{Q} reflects the trust level that could be given to the state-vector evolution model f . For example, the constant acceleration model is obviously incorrect for human motion leading to inconsistencies in the state-vector [35] and thus this should be reflected in \mathbf{Q} . Similarly, despite the bracelet form, AR markers have shown some sensitivity to occlusions. Fortunately, these are easily detectable. Thus, if the loss of all three markers data was detected, the corresponding elements of the EKF noise covariance matrix \mathbf{R} will be automatically updated to reflect the new state of input data. Consequently, two \mathbf{R} matrices were considered: the one corresponding to the whole markers trajectories when no occlusion of at least one marker has occurred and another one when all AR markers were occluded. In the latter, the missing markers poses were replaced by their previous values and their associated \mathbf{R} elements were given a large value ($1e^3$) indicating a lack of trust in these measurements. Considering the assumption that the parameters tuning is task-dependent [23], which is commonly used in EKF for motion analysis, the data of one randomly chosen subject was used to identify the elements of \mathbf{Q} . The same identified values were then used for the remaining subjects. Then, depending on the state-vector formulation, two methods to tune the parameters of \mathbf{R} and \mathbf{Q} were proposed in this study.

In the case 1, corresponding to the EKF-CA, an optimisation process aiming at tuning the parameters of the EKF-CA was used. The optimization process consists of finding $\mathbf{V} = [Q_\theta \ Q_{\dot{\theta}} \ Q_{\ddot{\theta}} \ R_p \ R_q \ R_a \ R_\Omega \ \mathbf{R}_\theta]^T$ (13×1) that minimizes the square difference between the estimated joint angles and the reference ones, obtained from the SS and MKO, respectively (see section IV). The optimization process is computationally costly. Thus, it is proposed to group some parameters and to set the elements corresponding to the constraints c2 to $1e^{-3}$. The choice of this value reflects the fact that the reference values are almost exact. The thirteen parameters that need to be identified are: three parameters Q_θ , $Q_{\dot{\theta}}$, and $Q_{\ddot{\theta}}$ corresponding to the joint angles, velocities and accelerations elements of the process covariance matrix \mathbf{Q} , and ten parameters corresponding to the measured positions R_p , quaternions R_q , linear accelerations R_a , angular velocities R_Ω , and to the six joints $\mathbf{R}_\theta = R_{\theta_i}$ in constraint c1 with $i = 1 \dots 6$ (assuming left and right symmetry). The problem of the optimal tuning of \mathbf{V} then boils down to:

$$\begin{aligned} \text{Find } \mathbf{V}^* \in \min_{\mathbf{V} \in \mathbb{R}^{13}} \sum_{j=1}^{N_J} \sum_{k=1}^N (\theta_{SS_{jk}} - \theta_{jk})^2 \\ \text{s.t. } \mathbf{V}^- \leq \mathbf{V} \leq \mathbf{V}^+ \end{aligned} \quad (9)$$

where N is the number of samples of the considered trial. \mathbf{V}^- and \mathbf{V}^+ represent the lower and upper boundaries, respectively. θ_{SS} (18×1) is the vector of the reference joint angles calculated using the SS (see Section IV). It should be noted that the gradient of this formulation is difficult to calculate since the equation in the EKF needs to be integrated.

Thus, Eq. 9 was solved using a local derivative-free solver proposed in the popular IPOPT library [34], leading to $\mathbf{V}_{opt} = [165.23 \ 3.18 \ 0.17 \ 0.13 \ 0.17 \ 1.09 \ 0.46 \ [0.09 \ 0.04 \ 0.03 \ 2.32 \ 0.1 \ 1]]^T$.

In the case 2 corresponding to EKF-FS, it is crucial to determine firstly the number of harmonics N_H required to accurately estimate the joint angles, velocities and accelerations. To do so, it is proposed to a priori solve the following fitting problem for different values of N_H :

$$\begin{aligned} \text{Find } \mathbf{a}^*, \mathbf{b}^*, \omega^* \in \min_{\mathbf{a}^*, \mathbf{b}^*, \omega^* \in \mathbb{R}^{2N_H + N_J + 1}} \sum_{j=1}^{N_J} \sum_{k=1}^N (\theta_{SS_{jk}} - \theta_{jk})^2 \\ \text{s.t. } 0 \leq \omega \end{aligned} \quad (10)$$

The Fourier representation is also used to obtain an analytical estimate of the joint velocities and accelerations. Thereafter, the results of this fitting process were used to determine a good trade-off between the number of parameters and the accuracy. Then, the parameter tuning is only based on a priori knowledge of the biomechanics of the task. The diagonal elements in \mathbf{Q} were set to a small value (0.01) to enforce their convergence to pseudo-constant values [35]. This guarantees the convergence of the corresponding error covariance \mathbf{P} to a value close to zero. The elements of \mathbf{R} related to the measurements were all empirically set to 0.1. As it regards the constraints (c1) related to the difference between estimated and average joint angles, the elements were set accordingly to the joint angle amplitude such that $\mathbf{R}_\theta = [1 \ 1 \ 1 \ 10 \ 1 \ 10]$ while the corresponding elements of the rigid body constraints (c2) were set to $1e^{-3}$.

III. EXPERIMENTAL VALIDATION

A. VIMU Prototype

The proposed motion capture system was developed to be affordable, user-friendly, compact and light-weight. VIMU prototypes are shown in Fig. 3.a. Each VIMU consists of one IMU (MPU-6050, Invensense) embedding a 3D gyroscope and a 3D accelerometer, and communicating at 100Hz using a Bluetooth wireless module. On the top of the IMU enclosure, measuring $4 \times 4 \times 2$ cm, three AR markers of 3.6×3.6 cm are mounted in a 3D printed bracelet form. The bracelet weights 40g and has different forms depending on the targeted segment. Its form is convenient to improve the AR markers detection rate during motions involving complex rotations. VIMUs were attached to human body segments using a velcro strap. Two standard 60Hz high-definition RGB USB cameras (USBFHD01, ELP) were used to track the AR markers poses. A custom program is developed in C++ to read all VIMUs data synchronously at a frequency rate of 60 Hz. The total cost of the proposed system was below 100 €.

B. Experimental protocol

Ten healthy male volunteers (age 25.4 ± 4.1 years old, weight 71.2 ± 16.0 Kg, height 1.7 ± 0.08 m) participated in the experiments. They gave their informed consent following

ethical regulation for non-invasive experiments of the University of Tokyo Agriculture and Technology (Tokyo, Japan) and of the University of Paris-Est Créteil (Créteil, France) where experiments took place. Once the VIMUs and retro-reflective markers were worn by the subject, at the different locations shown in Fig.2.b, every subject was told to stand statically, for less than one minute, in any comfortable posture that allows detection of VIMUs AR markers using the RGB cameras. Two RGB cameras were setup on tripods to be able to visualize the whole scene and detect all AR markers. The two RGB cameras parameters were obtained thanks to a standard calibration process based on a chArUcoBoard [36]. The data collected by the two cameras were thereafter expressed in the same coordinate system using a large AR marker located on the floor in the field of view of the two cameras (Fig.4). Meanwhile, anatomical landmarks of interest were pin-pointed using the affordable calibration wand and also captured using the SS (see Fig. 2.b). As a result, joint centres' positions, segments' lengths, as well as the local pose of each VIMU sensor relative to its corresponding segment anatomical frame, could be accurately calculated (Section. II.B). Hereafter, all subjects performed normal gait on a treadmill for one minute while wearing their running shoes. Each participant was free to choose his natural and comfortable gait speed.

C. Accuracy Assessment

The ability of the two proposed EKF state-vector formulations to accurately estimate the lower-limbs joint kinematics was assessed by calculating the RMSD and the Pearson Correlation Coefficient (CC) between the estimated and reference joint angles obtained by means of the SS. This choice is due to the fact that we were interested in estimating the temporal joints evolution, independently of the alignment and calibration offsets between the global reference frames of the different motion capture systems. Note that the sensor data are presented in international system of units while the outputs of the proposed system are given in degrees and centimeters since these units are more easily interpreted in a clinical context. The reference joint angles were obtained using a standard reference SS (6 Prime cameras, Optitrack) and a state of the art MKO tracking 22 reflective markers [24] with the multi-body model shown in Fig. 2. The joint angles computed with the SS follow the Conventional Gait Model [37]. Moreover, the knee flexion axis was defined orthogonal to the long axis of the femur and, at the hip, the joint coordinate system corresponded to 2-by-2 mutually orthogonal axes, similarly to the multibody model as shown in Fig.2.a. The data of the three VIMUs, located at the pelvis and both heels of the subjects as shown in Fig.2.b, were synchronously collected.

The gait deviation index (GDI) [3] was calculated using the estimated joint angles obtained from EKF-CA and from EKF-FS. Both results were compared with the mean GDI obtained from reference SS data, with a leave-one-subject-out cross validation. Estimates of nine subjects were used to build the basis required to estimate GDI of the 10th subject. A total of approximately 260 *healthy strides* were used to build the GDI basis matrix on 9 subjects. The strides of the 10th subject

were labelled as test strides, and used to calculate the GDI relatively to healthy strides. This operation was performed for each subject.

Moreover, based on the joint angles estimates and the forward kinematic model (see Eq. 1), the sacrum and heels positions were estimated and used to compute the gait stride lengths. The occurrences of heel strikes on the treadmill were first determined using the method developed in [38] and then used to estimate the corresponding stride lengths as in [39].

Finally, we have studied the differences between EKF-CA and EKF-FS results by verifying that they were statistically nonsignificant. This was done using the corresponding RMSD and CC of all estimated gait kinematics parameters and of the GDI through a paired t-test with an alpha level set a priori to 0.05. The Pearson Chi-square goodness-of-fit test was conducted to check for normal distribution in the data.

IV. RESULTS

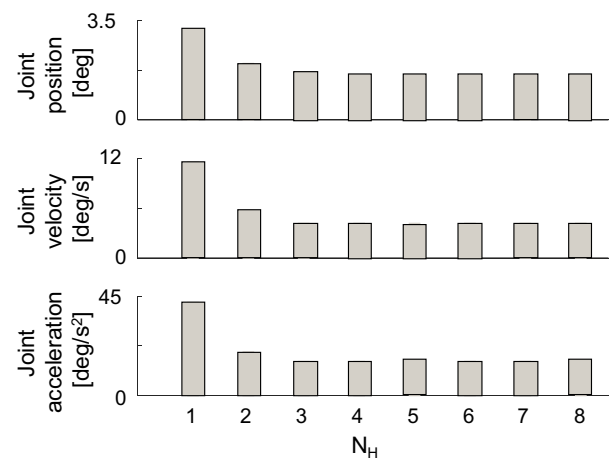


Fig. 5. Mean RMSD between the joint angles, velocities and accelerations estimated using Fourier series expansion and those obtained using the SS as a function of the number of harmonics N_H .

The mean RMSD of the joint angles, velocities and accelerations calculated using Fourier series representation and those estimated using the SS as a function of the number of harmonics N_H for a randomly selected subject are shown in Fig.5. To minimise the RMSD for the joint trajectories the number of harmonics was set to $N_H = 4$ accordingly to Fig. 5.

Fig. 6 shows the tracking performance using the proposed EKF-CA of the data collected by the VIMU sensor attached to the left heel while walking on the treadmill. Overall the input data were correctly tracked. The RMSD on the IMU data tracking were on average of $0.017 \text{ m}\cdot\text{s}^{-2}$ and of $0.029 \text{ rad}\cdot\text{s}^{-1}$ for the linear accelerations and for the angular velocities, respectively. For the AR marker positions and quaternions, the RMSD were on average of 3.8 cm and $2.2e^{-2}$, respectively. These relatively large tracking errors were due to the fact that the AR markers were not detected just after toe-off as exemplify by the zoomed windows of Fig.6. In fact at this moment the AR markers were positioned almost perpendicularly to the cameras. Despite these tracking errors, the EKF was able to

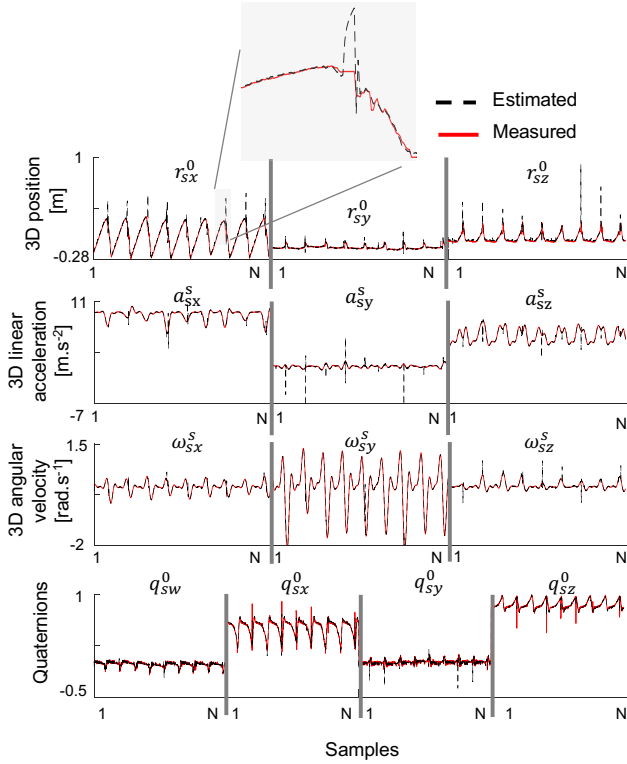


Fig. 6. Comparison between the input data of the VIMU attached to the left heel (red) and their tracking with the EKF-CA (dashed black). These are the 3D position and quaternion, 3D linear acceleration and 3D angular velocity. Note that the vertical grey lines are used to separate data that are presented for all N samples on the same subplot.

accurately estimate the eighteen joint angles of the lower-limbs model described in Fig. 2. This is valid for both the CA and FS state-vector formulations using the optimal parameters tuning. Fig. 7 exemplifies the accurate estimates of the 6 DoF that were used to calculate the pose of the pelvis with respect to R_0 for both approaches. In this figure, the corresponding average RMSD of the 3D pelvis position was equal to $0.4 \pm 0.0\text{cm}$ (CC 0.86 ± 0.20) and $1.6 \pm 0.9\text{deg}$ (CC 0.75 ± 0.16) in the case of EKF-CA and $0.4 \pm 0.0\text{cm}$ (CC 0.80 ± 0.30) and $2.1 \pm 0.7\text{deg}$ (CC 0.76 ± 0.21) in the case of EKF-FS. Fig. 8 shows a representative comparison of the left leg joint angles estimate obtained using both EKF-CA and EKF-FS approaches. The corresponding joint angles were satisfactorily estimated with an average RMSDs and CCs of $3.3 \pm 1.6\text{deg}$ and 0.76 ± 0.24 using the EKF-CA and of $3.3 \pm 1.2\text{deg}$ and 0.74 ± 0.20 using the EKF-FS, respectively.

Table I summarizes the results of the comparison between the lower-limb joint trajectories obtained with the proposed system and the SS, for each considered measure, based on the data from all subjects. Note that the results for the leg joints (i.e., hip, knee and ankle) were obtained over the average of left and right leg values. In general, results were similar between both EKF-CA and EKF-FS approaches. For the pelvis the average RMSDs were of $0.6 \pm 0.1\text{cm}$ (CC 0.84 ± 0.21) and $2.0 \pm 0.25\text{deg}$ (CC 0.60 ± 0.09) when using the EKF-CA and $0.7 \pm 0.0\text{cm}$ (CC 0.82 ± 0.22) and $2.2 \pm 0.4\text{deg}$ (CC 0.63 ± 0.18) when using the EKF-FS. For the leg joint angles, i.e., the hip, knee, and

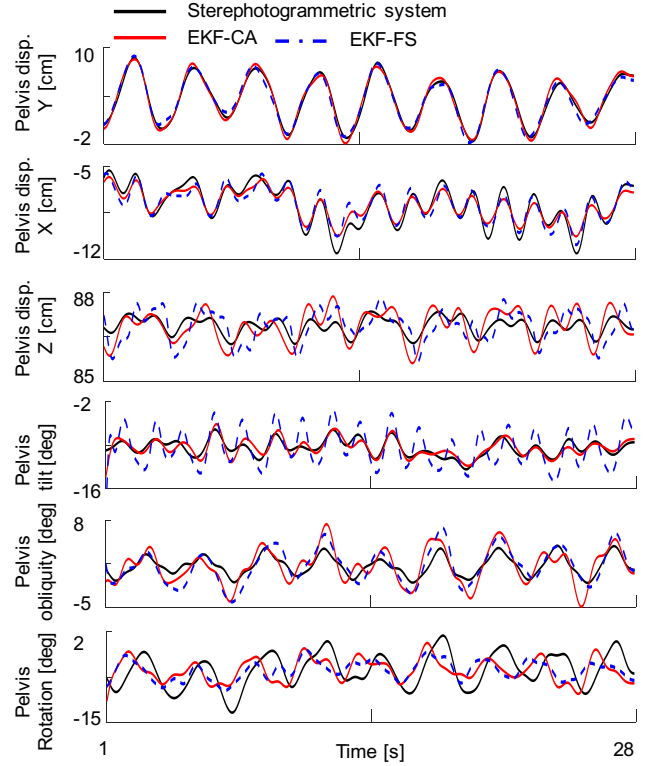


Fig. 7. Representative comparison of the pelvis joint trajectories obtained for a random selected subject while walking on a treadmill using the stereophotogrammetric system (black) and the proposed approach based either on a constant acceleration model (red) or on Fourier series expansion (dashed blue).

ankle joint angles, the average RMSDs for all subjects were equal to $3.9 \pm 1.0\text{deg}$ (CC 0.74 ± 0.19) and $3.9 \pm 0.8\text{deg}$ (CC 0.71 ± 0.19) when using EKF-CA and EKF-FS, respectively.

Table. II summarizes the mean and standard deviation RMSDs for right and left stride lengths estimates. These were calculated first using the VIMUs raw data (without processing) obtained from the AR markers pose estimation, then using both EKF-CA and EKF-FS estimated data, and were compared with respect to the SS. In both cases, the use of the EKF reduced the RMSD in stride length estimates. Specifically, the EKF-FS showed an improvement by 50% approximately passing from 2.6cm to 1.3cm.

Moreover, using the joint angles estimates for the detected strides, the GDI was also calculated. Results presented in Table II show that all the mean of GDIs (the reference and the estimated ones), are very close to 100 indicating a normal walking pattern.

The paired t-test results for RMSD and CC of each joint trajectory estimated using both the EKF-CA and EKF-FS implementations are reported in Table I. For all joints, results showed nonsignificant statistical differences ($p > 0.05$) between EKF-CA and EKF-FS with the exception of the pelvic tilt angle ($p_{RMSD} = 0.001$, $p_{CC} = 0.039$). Nevertheless, considering the above results, the gait kinematics estimates do not seem to be deteriorated by the poor estimate of the pelvic tilt angle. Errors in pelvis tilt angle also mechanically affects hip flexion-extension. However, while pelvis tilt is of

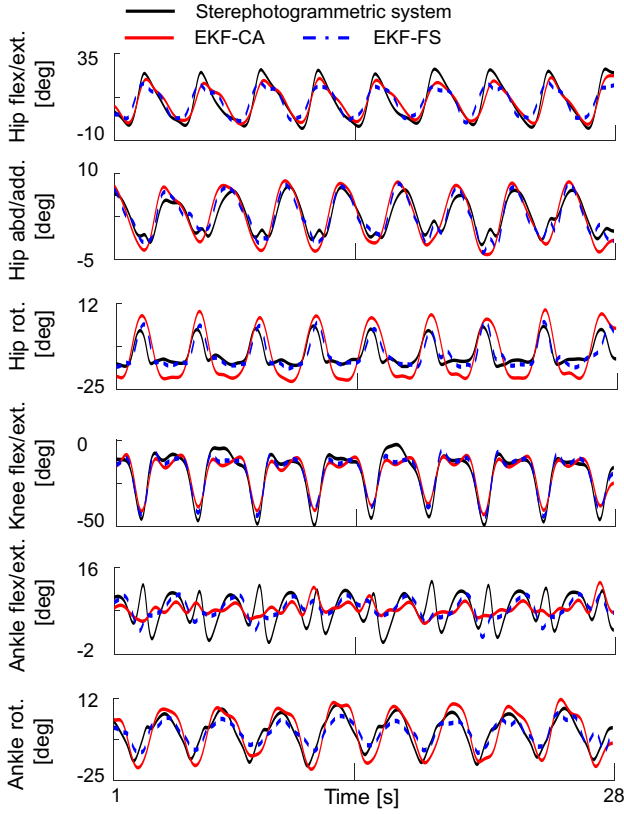


Fig. 8. Representative comparison of the left leg joint angles obtained for a random selected subject while walking on a treadmill using the stereophotogrammetric system (black) and the proposed approach based either on a constant acceleration model (red) or on Fourier series expansion (dashed blue)

limited amplitude during gait, hip flexion-extension is of large amplitude. Therefore, the segment angle errors do not translate into significant joint angle errors between EKF-CA and EKF-FS ($p = 0.200$ for RMSD and $p = 0.071$ for CC). A similar statistical test was performed in Table III for the RMSD of the stride lengths and for the GDI calculation. Both stride lengths' RMSD and GDI showed nonsignificant differences between EKF-CA and EKF-FS. The difference in GDI values resulting from each EKF implementation and SS data was also nonsignificant. On the other hand, the raw stride length values compared to their estimates based on the two EKFs (see Table III) revealed, as expected, significant differences. This emphasizes the improvement provided by the EKFs in either implementation.

Fig. 9 shows the evolution of the nine Fourier series coefficients corresponding to the ankle flexion angle estimated using EKF-FS. The coefficients converge toward pseudo-constant values showing, as expected, the gait periodicity.

V. DISCUSSION

The proposed system proposes to fuse visual and inertial data of three VIMUs attached to the pelvis, left and right heels into an EKF and a lower-limb kinematic model. This allowed handling the inaccuracies that result from using affordable sensors. Visual occlusion was greatly limited thanks to the

TABLE I
RESULTS OF THE COMPARISON BETWEEN EIGHTEEN LOWER-LIMBS JOINT TRAJECTORIES OBTAINED USING THE PROPOSED AFFORDABLE SYSTEM AND THE SS. RESULTS HAVE BEEN REPORTED AS MEAN \pm SD OVER ALL SUBJECTS AND RESULTS OF HIP, KNEE, AND ANKLE JOINTS HAVE BEEN OBTAINED OVER THE AVERAGE OF LEFT AND RIGHT LEG VALUES. RESULTS OF THE PAIRED T-TEST BETWEEN EKF-CA AND EKF-FS FOR RMSD AND CC OF EVERY JOINT TRAJECTORY HAVE ALSO BEEN REPORTED.

	Constant acceleration model assumption		Fourier series model assumption		p -Values	
	RMSD [cm]	CC	RMSD [cm]	CC	RMSD	CC
Pelvic Translations						
Translation Y θ_1	0.7 \pm 0.2	0.96 \pm 0.03	0.7 \pm 0.2	0.94 \pm 0.05	0.444	0.608
Translation X θ_2	0.7 \pm 0.1	0.96 \pm 0.02	0.8 \pm 0.1	0.96 \pm 0.03	0.472	0.273
Translation Z θ_3	0.5 \pm 0.2	0.59 \pm 0.19	0.7 \pm 0.2	0.56 \pm 0.25	0.182	0.796
Mean	0.6 \pm 0.1	0.84 \pm 0.21	0.7 \pm 0.0	0.82 \pm 0.22		
Joint Angles	RMSD [deg]	CC	RMSD [deg]	CC	RMSD	CC
Pelvis tilt	1.7 \pm 0.8	0.54 \pm 0.26	2.4 \pm 0.7	0.43 \pm 0.22	0.001	0.039
Pelvis obliquity	2.0 \pm 0.8	0.56 \pm 0.4	1.7 \pm 0.6	0.77 \pm 0.13	0.372	0.136
Pelvis rotation	2.2 \pm 0.8	0.70 \pm 0.19	2.4 \pm 0.7	0.69 \pm 0.29	0.610	0.382
Hip Flex./Ext.	3.6 \pm 0.9	0.95 \pm 0.02	3.9 \pm 1.1	0.93 \pm 0.02	0.200	0.071
Hip Abd./Add.	2.6 \pm 1.3	0.64 \pm 0.18	2.7 \pm 1.2	0.57 \pm 0.25	0.583	0.205
Hip Int./Ext.	4.1 \pm 1.5	0.79 \pm 0.20	4.2 \pm 1.1	0.73 \pm 0.25	0.872	0.421
Knee Flex./Ext.	5.1 \pm 1.1	0.91 \pm 0.06	4.8 \pm 0.9	0.92 \pm 0.04	0.152	0.127
Ankle Flex./Ext.	3.1 \pm 0.8	0.42 \pm 0.27	3.1 \pm 1.1	0.45 \pm 0.25	0.841	0.941
Ankle Inv./Eve.	5.0 \pm 1.6	0.72 \pm 0.20	4.7 \pm 1.5	0.64 \pm 0.25	0.516	0.085
Mean	3.5 \pm 1.2	0.71 \pm 0.17	3.5 \pm 1.0	0.69 \pm 0.18		

TABLE II
COMPARISON BETWEEN USING THE AFFORDABLE RAW/SS DATA AND THE DATA ESTIMATED BY EKF, FOR OBTAINING THE STRIDE LENGTHS AND GAIT DEVIATION INDEX (GDI), OVER ALL SUBJECTS AND STRIDES.

Stride lengths	Raw data		EKF-CA		EKF-FS	
	RMSD [cm]	NRMSD [%]	RMSD [cm]	NRMSD [%]	RMSD [cm]	NRMSD [%]
Right	2.6 \pm 1.1	3.8 \pm 1.6	1.7 \pm 0.0	2.3 \pm 0.5	1.4 \pm 0.0	1.9 \pm 0.2
Left	2.6 \pm 1.0	3.8 \pm 1.5	1.6 \pm 0.0	2.2 \pm 0.8	1.3 \pm 0.0	1.7 \pm 0.5
	SS data	EKF-CA	EKF-FS			
GDI	99.66 \pm 3.87	98.44 \pm 7.82	98.38 \pm 7.97			

TABLE III
THE p -VALUES RESULTING FROM THE PAIRED T-TEST BETWEEN RAW/SS DATA, EKF-CA, AND EKF-FS FOR BOTH RMSD OF STRIDE LENGTHS AND GDI.

p -Values			
Stride lengths RMSD	EKF-CA / Raw data	EKF-FS / Raw data	EKF-CA / EKF-FS
Right	0.049	0.013	0.093
Left	0.016	0.005	0.183
GDI	EKF-CA / SS data	EKF-FS / SS data	EKF-CA / EKF-FS
	0.648	0.280	0.984

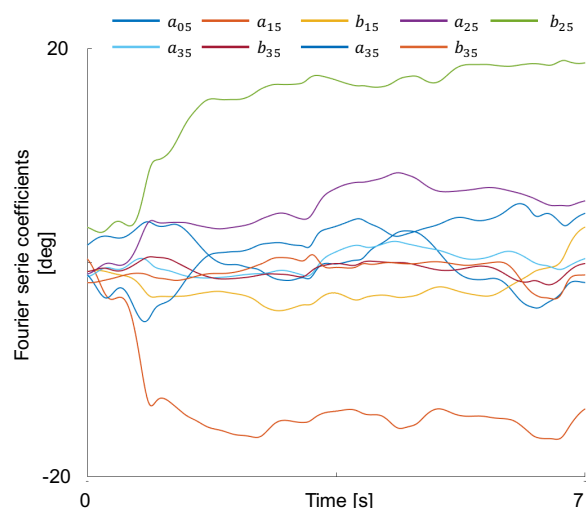


Fig. 9. Representative evolution of the Fourier series coefficients corresponding to the ankle rotation joint (θ_{12}) estimated by the EKF-FS.

bracelet form of the VIMUs. Rigid body constraints were integrated in the EKF formulation allowing to reconstruct marker loss based on its non-occluded partners within the same VIMU. This was possible as the rigid transformation between each two VIMU markers was priory calibrated (section II.B.1). Nevertheless, depending on the motion phases, the complete occlusion of the three AR markers could not be entirely avoided. Yet, the EKF was able to cope with such occlusion by relying more on IMU data thanks to the online update of the measurement covariance matrix \mathbf{R} . This is shown in the zoomed capture of Fig. 6 where one can note that the EKF is not tracking inconsistent visual data. However, it is important to note that the occlusion was never longer than few seconds. We have shown in a recent study that EKF approach was able to cope efficiently with such short occlusion [22]. Thus we can conclude that the proposed method is not very sensitive to occlusion with the current setup.

Moreover, constraints having the effects of joint springs pushing the joint angles toward mean values were added to the EKF as soft constraints in the measurement model \mathbf{h} . This was motivated by the fact that a sparse placement of the sensors was used. Thus, for the same VIMUs tracking, different joint configurations, possibly within the joint limits, are possible. Basically one could imagine to have springs of different stiffness attached to each joint. This concept is similar to the popular potential field approach [40] commonly used within the robotics community.

The proposed approach has been validated with a direct comparison to the joint kinematics estimated using a reference SS. For consistency in the joint trajectories comparison, the SS data were also processed with MKO using the same multi-body model. When tracking the same data, EKF and MKO are expected to provide very similar results [24]. Over the fifteen lower-limbs joint angles the mean RMSD was of 3.5deg. These results show the ability of the proposed approach to handle the 3D IMU drift without magnetometers involvement while compensating a sparse sensors placement.

For both EKF implementations, the knee and hip flexion-extension angles estimates presented a relatively high CC of 0.927 on average and a low RMSD ($\leq 5\text{deg}$). In contrast, the CCs of the ankle dorsi/plantar flexion (0.435) and the pelvic tilt angle (0.485) estimates are the poorest but with a satisfactory RMSD of 3.1deg and 2.0deg, respectively. Moreover the p -value of the pelvic tilt angle is the only angle displaying significant different results between the two EKF implementations. This might be explained by the small angle amplitude of this joint while walking on the treadmill. Moreover, the tracking error is not equally distributed among all joint trajectories due to constraints $c1$ (similar to joint springs) and to different parameters values in the covariance matrix \mathbf{R} .

When compared to reference kinematics (bone pins or fluoroscopy), a systematic review reported errors between 1 and 22deg at the knee joint when tracking retro-reflective markers placed on both thigh and shank segments [24]. The proposed method outperforms related studies dealing with a sparse sensor placement [8] (RMSD of 10deg) or when using a simple visual observation (RMSD of 9deg [10]) that is often still used by physiotherapists (see section I.A for further details on these studies). When using a single IMU per investigated segment and a quasi-periodic assumption it is possible to reach a better accuracy without monitoring the ankle joints (RMSD of 2.4deg) [10]. Of course this increases the complexity and the price of the overall system.

While similar results were observed by comparing the use of a constant acceleration model and a Fourier series expansion in the EKF (see Table. I), it is important to note that pre-processing to obtain these results is very different with the two methods. In the former, the model assumption is inherently inaccurate in human motion. The literature related to the use of EKF for motion analysis rarely proposes methods to model these inaccuracies and to reflect them in the parameters tuning. Cerveri et al. [41] were the sole authors to propose a frequency based approach but this requires to have a prior knowledge of the joint angle evolution, applies to cyclic motions and was not tested with low-cost sensors data. We have shown in a previous study [16], using a RGB-D camera and a markerless joint center estimate algorithm, that an optimization process outperforms frequency based method and is absolutely required to tune the EKF parameters. Nevertheless, when using markerless data as input and despite an optimal tuning of the EKF parameters and a simplified experimental setup [16], it was not possible to achieve a joint estimate RMSD lower than 9.7deg on average when compared to a reference SS. Specifically, for tasks closely related to walking, a RMSD up to 20deg was observed for the hip joints. This emphasizes that markerless approaches are not yet accurate enough to estimate human motion, and the superiority of our proposed system being a good and affordable compromise between reference SS and markerless systems.

Unfortunately, an optimal parameters tuning approach is time consuming, since it requires to run the EKF several thousands of times leading to at least 10h of calculation time. Moreover, it can not be performed independently from a SS. Although this optimisation is task dependant and can be performed once

for different subjects, it requires that one subject performs the task wearing simultaneously VIMUs and retro-reflective markers at least once. In the latter, when using a Fourier series model, the EKF parameters can be directly set based on a priori knowledge. As exemplified in Fig. 9, the Fourier series converge towards pseudo-constant values, confirming, as expected, the quasi-periodicity of the joint trajectories. The quasi-periodicity of gait was already used for motion analysis in the literature [10]–[12] and seems to be a valid assumption even for patients.

Regarding values obtained by the optimization process for the tuning of the EKF-CA, they are difficult to be interpreted as the elements of \mathbf{Q} and \mathbf{R} are interdependent. However we can observe that for the noise covariance matrix they are of the same order of magnitude than the ones set empirically for the EKF-FS and that for the joint soft constraints the ratio of one for ten between some angles is also respected.

Despite the differences in the two EKF implementations, the paired t-test results obtained for GDI and most gait kinematics (see Tables I and III) were not statistically significant. This suggests that both EKF-CA and EKF-FS can be used interchangeably for gait kinematics estimation while emphasizing the relevance of using an optimal tuning of the EKF parameters and/or a realistic process model that describes the joints evolution during movement.

The position of the joint centers were estimated thanks to an anatomical calibration even if functional calibration movements could have been used [9]. However, the proposed system targets patients that might not be able to perform exciting postures or motions [9]. Nevertheless, for patients having a certain level of recovery, such as patients at home, the fact that with the proposed system position and orientation data could be directly provided by the AR markers would make the estimate of joint centers straightforward [9]. Note that in both EKF formulations, the difference in joint angles estimates might be explained by the constant body segment lengths, computed statically as the distance between subsequent joint centre positions, as well as the VIMU-to-body-segment translational and rotational offsets. The latter may occur due to inaccuracies in AR markers measurements at initial static calibration, soft tissue artifacts, and the assumption of sensors-to-body-segments rigid connection. In such case, both segments lengths and VIMUs local poses could be modelled in the EKF but at the cost of increasing the state-vector dimensions.

The NRMSD of the strides lengths estimated with the SS and estimated when using the corrected EKF data was on average 1.5cm or 2% for both implementation cases (see Table II). This result is better than the 3cm reported in a recent study when using IMU data collected at both heel levels [42]. Many studies have intended to estimate similar spatial gait parameters. For instance, Kose et al. reported a maximal error up to 3% for step length estimate [43]. However, their approach, despite its ease of use, suffers from signal features dependency limiting its scope to normal walking. It relies directly on the IMU signal for heel strike detection and assumes a support leg in full extension. The approach proposed by Ahmed et al. uses a camera but only at the head level and thus it suffers from the same problems [21]. These authors

do not directly compute stride or step length, however their foot trajectory estimates, even if acquired over long travelled distance, showed a relatively large RMSD varying from 5 to 13cm. A recent study compared different spatio-temporal gait parameters between healthy and amputee patients [44]. The patient's average stride length was about 1.05 m, while the normal walking was on average of 1.24 m. This difference is much larger than the accuracy provided by the proposed system. Thus, it would be easy to distinguish between the different walking modes.

Having correctly estimated joint trajectories and stride lengths allow to estimate other indexes such as the GDI. The GDI difference between each EKF and SS data was not statistically significant. Moreover, GDI results were all very close to 100. By definition, a GDI greater or equal to 100 means that the test strides are close enough to the healthy strides to be classified as normal ones. The obtained GDI illustrates that the joint trajectories, as well as stride lengths, estimated using the proposed approach have the potential to provide clinically useful indicators. In addition, from the EKF results, the evolution of the Fourier coefficients could be further used to evaluate the gait variability of patients [11].

The proposed system is composed of two RGB cameras. This number could be easily increased to match a regular motion capture system and target more complex tasks where the subject will not be facing the same set of cameras. Note that unlike classical stereophotogrammetric approaches, the proposed one is less sensitive to visual occlusion thanks to the simultaneous use of IMU data. However, as previously demonstrated [22], the subject should not be too long away from the camera(s) field of view as the IMUs drift will impact the accuracy of the joint angle estimate. This is inherent to any purely IMU based system [10]. Another limitation of purely IMU based system is that the absolute position of the body with respect to the ground, and for instance to a forceplate, remains undetermined which impedes joint kinetics to be computed. The addition of externally fixed RGB camera(s) solves this issue. This way, the proposed affordable system shall be suitable for both kinematics and kinetics. The proposed system could be used outside of the laboratory but it will still require trained staff to perform the anatomical calibration. It could be used to monitor gait in small clinical centers or at doctor or physiotherapist practice for example. This is already the case of much more expensive devices that can monitor solely the knee joint angle [45].

VI. CONCLUSION

A new approach for motion analysis based on the use of VIMU sensors and an EKF has been proposed and validated in this paper. VIMU sensors can be considered a good compromise between, on the one hand, SS systems which are costly and cumbersome and, on the other hand, purely IMU based methods which requires one IMU per segment and advanced anatomical calibrations to reach an accuracy below 5deg. On the computational side, the propose EKF-FS represent one of the most advance implementations of this approach (including constraints and a clear rationale for the filter parameter tuning).

ACKNOWLEDGMENT

We thank all subjects of the GVLAB for their participation in the experiments as well as Patrick Karasinski and Florian Huet for their help in developing the VIMU prototype.

REFERENCES

- [1] R. Baker, "Gait analysis methods in rehabilitation", *J. NeuroEngineering Rehabil.* 3, 2006.
- [2] M.G. Benedetti, et al., "SIAMOC position paper on gait analysis in clinical practice: General requirements, methods and appropriateness. Results of an Italian consensus conference", *Gait Posture*, pp. 252 - 260, 2017.
- [3] M.H. Schwartz, and A. Rozumalski, "The Gait Deviation Index: a new comprehensive index of gait pathology", *Gait Posture*, pp. 351-35, 2008.
- [4] A. Cappozzo, "Minimum measured-input models for assessment of motor ability", *J. of biomech.* pp. 437-46, 2002.
- [5] A. Ferrari, P. Ginis, M. Hardegger, F. Casamassima, L. Rocchi and L. Chiari, "A Mobile Kalman-Filter Based Solution for the Real-Time Estimation of Spatio-Temporal Gait Parameters", *IEEE T. Neur. Sys. Reh.*, pp. 764 - 773, 2015.
- [6] A.M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing", *IEEE T. Neur. Sys. Reh.*, pp. 1346-1356, 2006.
- [7] M. El-Gohary, and J. McNames, "Human joint angle estimation with inertial sensors and validation with a robot arm", *IEEE. Trans. Biomed. Eng.*, pp. 1759-1767, 2015.
- [8] L.W. Sy, M. Raitor, M. Del Rosario, H. Khamis, L. Kark, N.H. Lovell, S. Redmond, "Estimating Lower Limb Kinematics using a Reduced Wearable Sensor Count", *IEEE T. Neur. Sys. Reh.*, doi: 10.1109/TBME.2020.3026464, 2020.
- [9] T. Seel, J. Raisch, and T. Schauer, "IMU-Based Joint Angle Measurement for Gait Analysis", *Sensors*, pp. 6891-6909, 2014.
- [10] V. Joukov, V. Bonnet, M. Karg, G. Venture, D. Kulic, "Rhythmic Extended Kalman Filter for Gait Rehabilitation Motion Estimation and Segmentation", *IEEE T. Neur. Sys. Reh.*, pp. 764-773, 2015.
- [11] D.J. Tsang, M. Lukac, A.E. Martin, "Characterization of statistical persistence in joint angle variation during walking", *Hum. Mov. Sci.*, vol. 68, 102528 2019.
- [12] J. Ziegler, A. Reiter, H. Gattringer, A. Müller, "Simultaneous identification of human body model parameters and gait trajectory from 3D motion capture data", *Med. Eng. Phy.*, pp. 193-202, 2020.
- [13] A.M.A. Sabatini, G. Ligorio, A. Mannini, "Fourier-based integration of quasi-periodic gait accelerations for drift-free displacement estimation using inertial sensors", *BioMed. Eng. Online* 14, 2015.
- [14] S. Chakraborty, A. Nandy, T. Yamaguchi, V. Bonnet, G. Venture, "Accuracy of image data stream of a markerless motion capture system in determining the local dynamic stability and joint kinematics of human gait", *J. Biomechanics*, 109718, 2020.
- [15] L. Zongmian, S. Jiri, J. Carpentier, I. Laptev, N. Mansard, and J. Sivic, "Estimating 3D Motion and Forces of Person-Object Interactions From Monocular Video", *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR'19)*, pp. 8640-8649, 2019.
- [16] J. Colombel, V. Bonnet, D. Daney, R. Dumas, A. Seilles, and F. Charpillet, "Physically Consistent Whole-Body Kinematics Assessment Based on an RGB-D Sensor. Application to Simple Rehabilitation Exercises", *Sensors*, 20, 2020.
- [17] B. Bonnechère, V. Sholukha, L. Omelina, S. Van Sint, and B. Jansen, "3D Analysis of Upper Limbs Motion during Rehabilitation Exercises Using the KinectTM Sensor: Development, Laboratory Validation and Clinical Application", *Sensors*, 2018.
- [18] G. Nagymate, R.M. Kiss, "Affordable gait analysis using augmented reality markers", *Plos One*, 14, 2019.
- [19] C. Chen and J. Roozbeh and N. Kehtarnavaz, "A survey of depth and inertial sensor fusion for human action recognition", *Multimedia Tools and Applications*, pp. 4405-4425, 2017.
- [20] S. Feng, and R. Murray-Smith, "Fusing Kinect sensor and inertial sensors with multi-rate Kalman filter", *IET Conference on Data Fusion Target Tracking 2014: Algorithms and Applications*, pp. 1-8, 2014.
- [21] Ahmed, and S. Roumeliotis, "A Visual-Inertial Approach to Human Gait Estimation", in *IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 1-8, Brisbane, Australia, 2018.
- [22] R. Mallat, V. Bonnet, M. Khalil, S. Mohammed, "Upper Limbs Kinematics Estimation Using Affordable Visual-Inertial Sensors", *IEEE Trans. Autom. Sci. Eng.*, 2020 (accepted for publication).
- [23] F. De Groote, T. De Laet, I. Jonkers, J. De Schutter, "Kalman smoothing improves the estimation of joint kinematics and kinetics in marker-based human gait analysis", *J. Biomech.* pp. 3390-3398, 2008.
- [24] M. Begon, M.S. Andersen and R. Dumas, "Multibody kinematics optimization for the estimation of upper and lower limb human joint kinematics: a systematized methodological review", *J. Biomech. Eng.*, 2018.
- [25] J. Diebel, "Representing attitude: Euler angles, unit quaternions, and rotation vectors", *Matrix*, vol. 58, no. 15-16, pp. 15, 2006.
- [26] A. Cappozzo, F. Catani, U. Della Croce, A. Leardini, "Position and orientation in space of bones during movement: anatomical frame definition and determination", *Clin Biomech.* pp. 171-188, 1995.
- [27] D. Tedaldi, A. Pretto, and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments", *IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 3042-3049, Hong Kong, China, 2014.
- [28] I. Soderkvist, P.A. Wedin, "Determining the movements of the skeleton using well-configured markers", *J. Biomech.*, pp. 1473-1477, 1993.
- [29] A.L. Bell, R.A. Brand, D. R. Pedersen, "Prediction of hip joint centre location from external landmarks", *Hum. Mov. Sci.* pp. 3-16, 1989.
- [30] J. Rose, J.G. Gamble, and J.M. Adams "Human Walking", *Philadelphia: Lippincott Williams and Wilkins*, 2006.
- [31] H.K. Ramakrishnan, and M.P. Kadaba, "On the estimation of joint kinematics during gait", *J. Biomech.*, vol. 24, no. 10, pp. 969-977, 1991.
- [32] A.Y. Aravkin, et al., "Optimization viewpoint on Kalman smoothing with applications to robust and sparse estimation", in *Compressed sensing & sparse filtering*, pp. 237-280, 2014.
- [33] P. Zarchan and H. Musoff, "Fundamentals of Kalman Filtering: A Practical Approach", *American Institute of Aeronautics and Astronautics*, Virginia, 2000.
- [34] M. J. D. Powell, "Direct search algorithms for optimization calculations", *Acta Numerica* 7, pp. 287-336, 1998.
- [35] V. Bonnet, R. Dumas, A. Cappozzo, V. Joukov, et al., "A constrained extended Kalman filter for the optimal estimate of kinematics and kinetics of a sagittal symmetric exercise", *J. Biomech.*, pp.140-147, 2017.
- [36] G. Bradski, "The OpenCV Library", *Dr. Dobb's Journal of Software Tools*, 2000.
- [37] F. Leboeuf, R. Baker, A. Barré, J. Reay, R. Jones, M. Sangeux, "The conventional gait model, an open-source implementation that reproduces the past but prepares for the future", *Gait and Posture*, vol. 69, pp. 235-241, 2019.
- [38] J.A. Zeni, J.G. Richards, and J.S. Higginson, "Two simple methods for determining gait events during treadmill and overground walking using kinematic data", *Gait posture*, vol. 27, pp. 710-714, 2008.
- [39] K. Supakkul, "Using positional heel-marker data to more accurately calculate stride length for treadmill walking: a step length approach", *arXiv:1710.09030*, 2017.
- [40] O. Khatib, "The potential field approach and operational space formulation in robot control", *Adaptive and learning systems*, pp. 367-377, 1986.
- [41] P. Cerveri, M. Rabuffetti, A. Pedotti, and G. Ferrigno, "Real-time human motion estimation using biomechanical models and non-linear state-space filters", *Med. Biol. Eng. Comput.*, vol. 41, no.2, pp. 109-123, 2003.
- [42] M. Hao, K. Chen and C. Fu, "Smoothing-Based 3-D Foot Trajectory Estimation Using Inertial Sensors", *IEEE T. Neur. Sys. Reh.*, pp. 3534-3542, 2019.
- [43] A. Kose, A. Cereatti, and U. Della Croce, "Bilateral step length estimation using a single inertial measurement unit attached to the pelvis", *J. Neuroeng. Rehabil.*, 9, 2012.
- [44] J. Kim, N. Colabianchi, J. Wensman, and D.H. Gates, "Wearable Sensors Quantify Mobility in People with Lower Limb Amputation during Daily Life", *IEEE T. Neur. Sys. Reh.*, pp. 1282-1291, 2020.
- [45] Z. Yu, Y. Zilong, W. Shaobai, H. Wenhan, M. Limin, H. Huayang, X. Hong, "The KneeKG system: a review of the literature", *Gait and Posture*, vol. 41, pp. 763-768, 2015.
- [46] F.J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded Up Detection of Squared Fiducial Markers", *Image Vis. Comput.*, 2018.