

# An optimization-based approach to human body motion capture using inertial sensors

Manon Kok\* Jeroen D. Hol\*\* Thomas B. Schön\*\*\*

\* *Division of Automatic Control, Linköping University, SE-581 83 Linköping, Sweden (email: manko@isy.liu.se)*

\*\* *Xsens Technologies B.V., P.O. Box 559, 7500 AN Enschede, the Netherlands (e-mail: jeroen.hol@xsens.com)*

\*\*\* *Department of Information Technology, Uppsala University, SE-751 05 Uppsala, Sweden (e-mail: thomas.schon@it.uu.se)*

---

**Abstract:** In inertial human motion capture, a multitude of body segments are equipped with inertial measurement units, consisting of 3D accelerometers, 3D gyroscopes and 3D magnetometers. Relative position and orientation estimates can be obtained using the inertial data together with a biomechanical model. In this work we present an optimization-based solution to magnetometer-free inertial motion capture. It allows for natural inclusion of biomechanical constraints, for handling of nonlinearities and for using all data in obtaining an estimate. As a proof-of-concept we apply our algorithm to a lower body configuration, illustrating that the estimates are drift-free and match the joint angles from an optical reference system.

Keywords: Human body motion capture, optimization, maximum a posteriori estimation, inertial sensors, 6D pose estimation.

---

## 1. INTRODUCTION

Human body motion capture is used for many applications such as character animation, sports and biomechanical analysis [Xsens Technologies B.V., 2013]. It focuses on simultaneously estimating the relative position and orientation of the different body segments (expressed in terms of the joint angles) and estimating the absolute position of the body. Motion capture is often performed using either vision-based technologies [Moeslund et al., 2006] or using inertial sensors. The main advantage of using inertial sensors over vision-based technologies is that they are not restricted in space and do not require line of sight visibility [Welch and Foxlin, 2002]. In inertial human body motion capture, the human body is equipped with inertial measurement units (IMUs), consisting of 3D accelerometers, 3D gyroscopes and 3D magnetometers as shown in Fig. 1. Each body segment's position and orientation (pose) can be estimated by integrating the gyroscope data and double integrating the accelerometer data in time and combining these integrated estimates with a biomechanical model. Inertial sensors are successfully used for full body motion capture in many applications [Xsens Technologies B.V., 2013, Roetenberg et al., 2013, Kang et al., 2011, Yun and Bachmann, 2006].

Inertial sensors inherently suffer from integration drift. When using inertial sensors for orientation estimation they are therefore generally combined with magnetome-



Fig. 1. Examples of inertial motion capture. Upper left: olympic and world champion speed skating Ireen Wüst wearing an inertial motion capture suit with 17 inertial sensors. Upper right: graphical representation of the estimated orientation and position of the body segments. Lower left and right: experiment showing that line of sight visibility is not necessary for inertial motion capture.

---

\* This work is supported by MC Impulse, a European Commission, FP7 research project, by CADICS, a Linnaeus Center funded by the Swedish Research Council (VR) and by BALANCE, a European Commission, FP7 research project.

ters. Magnetometer measurements, however, are known to cause problems in motion capture applications since the magnetic field measured at the different sensor locations is typically different [Luinge et al., 2007, Cooper et al., 2009, Favre et al., 2008]. Including information from biomechanical constraints, i.e. information about the body segments being rigidly connected, can eliminate the need of using magnetometer measurements. Incorporating these constraints, the sensor’s *relative* position and orientation become observable as long as the subject is not standing completely still [Hol, 2011]. Estimating joint angles using a pair of inertial sensors, where each sensor estimates its own orientation using an extended Kalman filter (EKF) [Yuan and Chen, 2013] is therefore computationally cheap, but valuable information from biomechanical constraints is lost. Existing approaches therefore include the biomechanical constraints like for instance in Luinge et al. [2007] where an EKF is run using only the accelerometer and gyroscope measurements and a least-squares filter is added to incorporate the biomechanical constraints.

To allow for natural inclusion of biomechanical constraints, we introduce a new optimization-based approach for inertial motion capture. Compared to filtering approaches, optimization-based approaches are computationally expensive. Recent developments in both computational power and in available algorithms have, however, opened up possibilities for solving large-scale problems efficiently and even in real-time [Mattingley and Boyd, 2010]. Using an optimization formulation of the problem, a smoothing estimate can be obtained and nonlinearities can be handled. It also opens up possibilities for simultaneously estimating calibration parameters and for incorporating non-Gaussian noise.

The paper is organized as follows. After introducing the problem formulation in Section 2, in Section 3 we will introduce the biomechanical model, discussing the relevant coordinate frames, variables and biomechanical constraints. In Section 4 we will subsequently introduce the dynamic and sensor models. In Section 6 we will discuss experimental results, focusing on a subproblem, namely a lower body configuration consisting of 7 sensors, assuming a known calibration and not including any position aiding. These experiments are intended to serve as a proof-of-concept. A more in-depth analysis including a comparison with other methods is planned for future work.

Note that using inertial sensors and biomechanical constraints only, the *absolute* position is not observable, i.e. any translation of the body’s position estimates will lead to an equally valid solution of the estimation problem. For example in the case of the speed skater in Fig. 1, the estimated pose of the speed skater will resemble the “true” motion, but the exact location on the ice rink is not observable. This unobservability typically results in a drift of the body’s absolute position over time. Because of this, it is not possible to compare our position estimates with those of the optical reference system and for now we focus on analysis of the joint angles. To estimate absolute position it is necessary to include e.g. GPS, ultra-wideband [Hol, 2011] or zero velocity updates when the foot is at stand still [Callmer, 2013, Woodman, 2010] and this is planned for future work.

## 2. PROBLEM FORMULATION

The use of inertial sensors for human body motion capture requires inertial sensors to be placed on different body segments. The knowledge about the placement of the sensors on the body segments and the body segments’ connections to each other by joints can be incorporated using a biomechanical model.

The problem of estimating the relative position and orientation of each body segment is formulated as a constrained estimation problem. Given  $N$  measurements  $y_{1:N} = \{y_1, \dots, y_N\}$ , a point estimate of the variables  $z$  can be obtained as a constrained maximum a posteriori (MAP) estimate, maximizing the posterior density function

$$\begin{aligned} \max_z \quad & p(z | y_{1:N}) \\ \text{s.t.} \quad & c_e(z) = 0, \end{aligned} \quad (1)$$

where  $c_e(z)$  represents the equality constraints. In our problem,  $z$  consists of both static parameters  $\theta$  and time-varying variables  $x_{1:N}$ . Using this together with the Markov property of the time-varying variables and the fact that the logarithm is a monotonic function, we can rewrite (1) as

$$\begin{aligned} \min_{z=\{x_{1:N}, \theta\}} \quad & \underbrace{-\log p(x_1 | y_1) - \log p(\theta)}_{\text{initialization}} \\ & \underbrace{-\sum_{t=2}^N \log p(x_t | x_{t-1}, \theta)}_{\text{dynamic model}} - \underbrace{\sum_{t=1}^N \log p(y_t | x_t, \theta)}_{\text{biomechanical/sensor model}} \\ \text{s.t.} \quad & c_{\text{bio}}(z) = 0. \end{aligned} \quad (2)$$

Obtaining the MAP estimate thus amounts to solving a constrained optimization problem where the constraints  $c_{\text{bio}}(z)$  originate from a biomechanical model. The cost function consists of different parts related to the initialization of the variables, a dynamic model for the time-varying states and a biomechanical and sensor model. More details about the variables, the different parts of the cost function and the constraints are provided in Sections 3 and 4.

The optimization problem (2) is solved using an infeasible start Gauss-Newton method [Boyd and Vandenberghe, 2004]. The number of variables in the problem will become large already for short experiments and a small number of segments. The problem (2) can, however, still be solved efficiently due to its inherent sparseness.

## 3. BIOMECHANICAL MODEL

A biomechanical model represents the human body as consisting of body segments connected by joints. In the example application in Fig. 1 the body is modeled as consisting of 23 segments, whereas Fig. 2 illustrates two of these body segments. These can be thought of as the upper and lower leg, each with a sensor attached to it. The main purpose of Fig. 2 is to introduce the different coordinate frames, variables and calibration parameters. These definitions can straightforwardly be extended to any sensor and any body segment. The relevant coordinate frames are:

**The local coordinate frame  $\mathbf{L}$**  aligned with the local gravity vector, with the  $z$ -axis pointing up. The

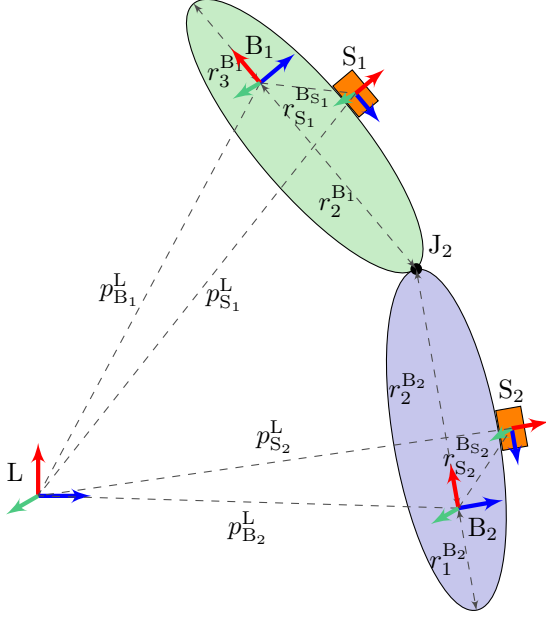


Fig. 2. Connection of two segments and definition of the variables and coordinate frames.

horizontal directions are defined according to any convenient choice of local coordinates.

**The body segment coordinate frame  $B_j$**  fixed to the bone in body segment  $B_j$ . Its origin can be anywhere along the bone, but it is usually in the center of rotation of a joint.

**The sensor coordinate frame  $S_i$**  of the moving IMU  $S_i$ . Its origin is located in the center of the accelerometer triad and its axes are aligned to the casing. All measurements of the IMU are resolved in this frame.

In setting up the optimization problem (2), the first step is to define the set of sensors  $\mathcal{S}$ , the set of body segments  $\mathcal{B}$  and the set of joints  $\mathcal{J}$  in the problem. Each inertial sensor needs to be mounted on the body, and sensor  $S_i$  is assumed to be placed on body segment  $B_{S_i}$ . The distance  $r_{S_i}^{B_{S_i}}$  and orientation  $q^{B_{S_i}S_i}$  of sensor  $S_i$  with respect to body segment  $B_{S_i}$  are without loss of generality assumed to be known from calibration.

Our knowledge of the human body can be used to identify which body segments are connected by which joints, i.e. the set  $\mathcal{B}_{J_k}$  needs to be determined for each joint  $J_k$ . To express the location of the joint in the body frames of the connected body segments, the distances  $r_k^{B_j}$  from the body frame  $B_j$  to joint  $k$ , need to be defined for all joints  $J_k \in \mathcal{J}$  and all  $B_j \in \mathcal{B}_{J,k}$ . We assume without loss of generality that they are known from calibration. Generally, all joints are assumed to be ball-and-socket joints, but we incorporate additional knowledge about a subset of the joints, denoted by  $\mathcal{H}$ , which we assume to be hinge joints.

For reasons that will be discussed in Section 4, we define the set of time steps in the optimization as  $\mathcal{T}$  rather than explicitly summing over all time steps  $t = 1 \dots N$  as in (2). The variables in the optimization problem are then given by

- the position  $p_{S_i,t}^L$  and velocity  $v_{S_i,t}^L$  of sensor  $S_i$  in the local frame  $L$ ,  $\forall S_i \in \mathcal{S}$  and  $\forall t \in \mathcal{T}$ ,

- the orientation  $q_t^{LS_i}$  of sensor  $S_i$  with respect to the local frame  $L$ ,  $\forall S_i \in \mathcal{S}$  and  $\forall t \in \mathcal{T}$ ,
- the position  $p_{B_j,t}^L$  of body segment  $B_j$  in the local frame  $L$ ,  $\forall B_j \in \mathcal{B}$  and  $\forall t \in \mathcal{T}$ ,
- the orientation  $q_t^{LB_j}$  of body segment  $B_j$  with respect to the local frame  $L$ ,  $\forall B_j \in \mathcal{B}$  and  $\forall t \in \mathcal{T}$ ,
- the gyroscope bias  $b_{\omega,S_i}$  of sensor  $S_i$ ,  $\forall S_i \in \mathcal{S}$ ,
- the mean acceleration state of one of the sensors  $S_i \in \mathcal{S}$ ,  $\forall t \in \mathcal{T}$ .

Defining the number of sensors as  $N_S$  and the number of body segments as  $N_B$ , the number of variables in the optimization problem is  $z \in \mathbb{R}^{(9N_S+6N_B+3)N+3N_S}$ . When we solve the optimization problem, we encode the rotation states using a three-dimensional state vector [Crassidis et al., 2007, Grisetti et al., 2010, Hol, 2011]. Throughout the paper, we typically interchangeably make use of the unit quaternion  $q^{LS}$  and the rotation matrix  $R^{LS}$  as representations of the orientation. The quaternion conjugate, representing the inverse rotation will be represented by  $(q^{LS})^c = q^{SL}$ . Similarly for the rotation matrix,  $(R^{LS})^T = R^{SL}$ .

More details about the gyroscope bias variables and the reason for the inclusion of the mean acceleration state will be given in Section 4.2.

Based on the biomechanical model it is possible to derive relations between the different variables. We will categorize them in three classes.

**Joints between the body segments.** The constraints  $c_{\text{bio}}(z)$  in the optimization problem (2) enforce the body segments to be connected at the joint locations at all times,

$$c_{\text{bio}}(z) = p_{B_m,t}^L + R_t^{LB_m} r_k^{B_m} - p_{B_n,t}^L - R_t^{LB_n} r_k^{B_n}, \quad (3)$$

$$B_n, B_m \in \mathcal{B}_{J_k},$$

which is included for all  $J_k \in \mathcal{J}$  and  $t \in \mathcal{T}$ . This leads to  $N_J$  constraints at each time step  $t$  in the optimization problem (2), where  $N_J$  is the number of joints.

**Placement of the sensors on the body segments.**

The position and orientation of sensor  $S_i$  can be expressed in terms of its position and orientation on body segment  $B_{S_i}$ . Ideally, this can be incorporated using equality constraints in (2). However, it is physically impossible to place the sensor directly on the bone. Hence, it has to be placed on the soft tissue and the sensor will inevitably move slightly with respect to the bone. We therefore model the position and orientation of sensor  $S_i$  on body segment  $B_{S_i}$  as

$$p_{S_i,t}^L = p_{B_{S_i},t}^L + R_t^{LB_{S_i}} \left( r_{S_i}^{B_{S_i}} + e_{p,t}^{B_{S_i}} \right), \quad (4a)$$

$$q_t^{LS_i} = q_t^{LB_{S_i}} q^{B_{S_i}S_i} \exp \left( \frac{1}{2} e_{q,t}^{S_i} \right), \quad (4b)$$

where we assume  $e_{p,t}^{B_{S_i}} \sim \mathcal{N}(0, \Sigma_p)$  and  $e_{q,t}^{S_i} \sim \mathcal{N}(0, \Sigma_q)$ .

**Rotational freedom of the joints.** For some joints, it is known that their rotation is (mainly) limited to one or two axes. An example of this is the knee which is a hinge joint, although it can in practice flex a little around the other axes too. Minimizing

$$e_{J_k,t} = \begin{bmatrix} n_1^\top \\ n_3^\top \end{bmatrix} \left( R_t^{\text{LB}_m} \right)^\top R_t^{\text{LB}_n} n_2, \quad B_n, B_m \in \mathcal{B}_{J_k}, \quad (5)$$

where  $n_1$ ,  $n_2$  and  $n_3$  denote the different axis directions and  $e_{J_k,t} \sim \mathcal{N}(0, \Sigma_k)$ , will minimize the rotation around any but the  $n_2$ -axis. This cost function can be included at any time  $t$  for any joint  $k$  that is a hinge joint, i.e.  $\forall J_k \in \mathcal{H}, \forall t \in \mathcal{T}$ . Note that inclusion of this knowledge is optional in the algorithm.

#### 4. DYNAMIC AND SENSOR MODELS

The sensor's position, velocity and orientation at each time instance can be related by a dynamic model in which the accelerometer and gyroscope measurements are used as inputs [Gustafsson, 2012, Hol, 2011]. In this work we choose a slightly different approach to reduce the number of variables in the optimization problem (2). To achieve high update rates using a relatively small number of variables, we use an approach similar to the one discussed by Savage [1998a,b]. Hence, strapdown inertial integration, in which the accelerometer and gyroscope signals are integrated, is run at high update rates. This leads to accelerometer measurements  $\Delta p$  and  $\Delta v$  representing a difference in position and velocity and gyroscope measurements  $\Delta q$  representing a difference in orientation. These are integrated for  $\frac{T_s}{T}$  times, where  $T_s$  is the sampling time of the inertial sensors and  $T$  is the sampling time used in the optimization problem (2).

##### 4.1 Dynamic model

The position, velocity and orientation of each sensor  $S_i$  are related from time  $t$  to time  $t+T$  using the accelerometer measurements  $\Delta p_t^{S_i}$ ,  $\Delta v_t^{S_i}$  and the gyroscope measurements  $\Delta q_t^{S_i}$ . The position and velocity states at each time step are modeled according to

$$p_{S_i,t+T}^L = p_{S_i,t}^L + T v_{S_i,t}^L + R_t^{\text{LS}_i} \left( \Delta p_t^{S_i} + w_{p,t}^{S_i} \right) + \frac{T^2}{2} g^L, \quad (6a)$$

$$v_{S_i,t+T}^L = v_{S_i,t}^L + R_t^{\text{LS}_i} \left( \Delta v_t^{S_i} + w_{v,t}^{S_i} \right) + T g^L, \quad (6b)$$

where  $\Delta p_t^{S_i}$  and  $\Delta v_t^{S_i}$  denote the inputs based on the accelerometer measurements. The noise terms are modeled as  $w_{p,t} \sim \mathcal{N}(0, Q_p)$  and  $w_{v,t} \sim \mathcal{N}(0, Q_v)$ . The earth gravity is denoted by  $g^L$ . The orientation states are modeled as

$$q_{t+T}^{\text{LS}_i} = q_t^{\text{LS}_i} \Delta q_t^{S_i} \exp \left( \frac{1}{2} w_{q,t}^{S_i} \right), \quad (6c)$$

where  $\Delta q_t^{S_i}$  denotes the gyroscope measurements, corrected for the estimated gyroscope bias, and  $w_{q,t}^{S_i} \sim \mathcal{N}(0, Q_q)$ .

Since (6) models the states in terms of their value at the previous time step, the state at the first time instance needs to be treated separately. The orientation  $q_1^{\text{LS}_i}$  of each sensor  $S_i$  is estimated using the first accelerometer and magnetometer sample of that sensor. Note that this is the only place in the algorithm where magnetometer measurements are used. The variables  $q_1^{\text{LS}_i}$  are then initialized around this estimated orientation with additive noise  $e_{q_1}^{S_i} \sim \mathcal{N}(0, \Sigma_{q_1})$ . The position  $p_{S_i,1}^L$  of one of the sensors

is without loss of generality initialized around zero with additive noise  $e_{p_1} \sim \mathcal{N}(0, \Sigma_{p_1})$ . This defines the origin of the local coordinate frame  $L$ .

##### 4.2 Sensor model

The gyroscope measurements are affected by a slowly time-varying sensor bias. For relatively short experiments, the sensor biases of all sensors  $S_i \in \mathcal{S}$  can be assumed to be constant. Hence, we include only one three-dimensional variable for each sensor to represent the gyroscope bias. This variable  $b_{\omega, S_i}$  is modeled as  $b_{\omega, S_i} \sim \mathcal{N}(0, \Sigma_{b_\omega})$ .

As described in Section 1, we do not include position aiding in our problem, resulting in only *relative* position and orientation observability. A problem that can be encountered for this case is that of so-called gravity leakage. Because the subject's *absolute* inclination is unobservable, the gravity vector risks being misinterpreted as an acceleration. In the case of stationary measurements, when the accelerometer only measures the gravity vector, the accelerometer measurements can be used as a source of absolute inclination information. In case of motion, the accelerometer measurements will measure an additional acceleration. It can, however, still be assumed that the mean acceleration over a certain time period is zero [Luinge, 2002]. We therefore assume that one sensor follows this acceleration model for all  $t \in \mathcal{T}$ , up to some noise  $e_a \sim \mathcal{N}(0, \Sigma_a)$ .

#### 5. RESULTING ALGORITHM

The biomechanical model from Section 3 and the dynamic and sensor models from Section 4 can be combined and used to describe the probability density functions in (2). Eliminating all constant terms from the optimization, this results in a constrained weighted least-squares problem. This problem is given by

$$\begin{aligned} \min_z \quad & \sum_{t \in \mathcal{T}} \sum_{S_i \in \mathcal{S}} \underbrace{\left( \|e_{p,t}^{S_i}\|_{\Sigma_p^{-1}}^2 + \|e_{q,t}^{B_{S_i}}\|_{\Sigma_q^{-1}}^2 \right)}_{\text{placement of sensors on body (4)}} \\ & + \underbrace{\left( \|w_{p,t}^{S_i}\|_{Q_p^{-1}}^2 + \|w_{v,t}^{S_i}\|_{Q_v^{-1}}^2 + \|w_{q,t}^{S_i}\|_{Q_q^{-1}}^2 \right)}_{\text{dynamic model (6)}} \\ & + \sum_{S_i \in \mathcal{S}} \underbrace{\|b_{\omega, S_i}\|_{\Sigma_{b_\omega}^{-1}}^2}_{\text{gyroscope bias}} + \sum_{t \in \mathcal{T}} \sum_{J_k \in \mathcal{H}} \underbrace{\|e_{J_k,t}\|_{\Sigma_k^{-1}}^2}_{\text{hinge (5)}} \\ & + \underbrace{\|e_{p_1}\|_{\Sigma_{p_1}^{-1}}^2 + \sum_{S_i \in \mathcal{S}} \|e_{q_1}^{S_i}\|_{\Sigma_{q_1}^{-1}}^2}_{\text{initialization}} + \sum_{t \in \mathcal{T}} \underbrace{\|e_{a,t}\|_{\Sigma_a^{-1}}^2}_{\text{acceleration model}}, \\ \text{s.t. } \quad & c_{\text{bio}}(z) = p_{B_m,t}^L + R_t^{\text{LB}_m} r_k^{B_m} - p_{B_n,t}^L - R_t^{\text{LB}_n} r_k^{B_n}, \\ & B_n, B_m \in \mathcal{B}_{J_k} \quad \forall J_k \in \mathcal{J}, \forall t \in \mathcal{T}, \end{aligned} \quad (7)$$

where the constraints are based on (3).

The complete algorithm is summarized in Algorithm 1. Note that in our current implementation the optimization is performed over the entire data set and the computations are therefore done offline. We plan to extend the approach to a moving horizon approach [Rao et al., 2001] to enable processing of longer data sets and to allow for online estimation.

The covariance matrices in (7) representing the sensor covariances are determined using Allan variance analysis [El-Sheimy et al., 2008]. The covariance matrices related to the placement of the sensors on the body, the hinge constraint and the acceleration model, do not represent any physical quantities and are chosen more or less ad hoc. Experiments have shown that the solution of the optimization problem is not very sensitive to the tuning of these values.

The optimization (7) is started using an initial estimate of the variables  $z_0$ . All variables are initialized at zero except for the orientations at the first time step, which are initialized around their estimated orientation, as described in Section 4.2. This is an infeasible solution, justifying the need for an infeasible start optimization algorithm.

---

**Algorithm 1** Inertial human motion capture

---

- (1) Define the set of sensors  $\mathcal{S}$ , the set of body segments  $\mathcal{B}$  and the set of joints  $\mathcal{J}$ . Mount the inertial sensors on the body and
    - (a) define for each sensor  $S_i \in \mathcal{S}$  on which body segment  $B_{S_i} \in \mathcal{B}$  it is placed. Calibrate the system to obtain the position  $r_{S_i}^{B_{S_i}}$  and orientation  $q^{B_{S_i}S_i}$  of each sensor  $S_i \in \mathcal{S}$  on body segment  $B_{S_i} \in \mathcal{B}$ ,
    - (b) define the set of body segments  $\mathcal{B}_{J_k}$  connected to each joint  $k$  for all  $J_k \in \mathcal{J}$ . Calibrate the system to obtain the distances  $r_k^{B_j}$  of each body segment coordinate frames  $B_j \in \mathcal{B}_{J_k}$  to the different joints  $k$ ,
    - (c) define the subset  $\mathcal{H}$  of joints that are restricted in their rotations and can be regarded as a hinge joint.
  - (2) Perform an experiment collecting inertial measurements  $\Delta p_t^{S_i}, \Delta v_t^{S_i}$  and  $\Delta q_t^{S_i}$  and a magnetometer measurement at  $t = 1, y_{m,1}^{S_i}$ .
  - (3) Postprocess the data
    - (a) Initialize  $z_0$  and set  $l = 0$ .
    - (b) Determine the values of the cost functions and the constraints in (2), their Jacobians and the approximate Hessian of the cost function. Determine a step direction using an infeasible start Gauss-Newton algorithm and update  $z_l \rightarrow z_{l+1}$ .
    - (c) Set  $l := l + 1$  and iterate from 3(b) until the algorithm is converged and the solution  $z_{l+1}$  is feasible.
- 

## 6. EXPERIMENTS

We validated our approach with experiments using an MVN Awinda system [Xsens Technologies B.V., 2013] which is a wireless inertial motion capture system with 17 sensors attached to different body segments as shown in Fig. 3. An optical motion capture system has been used as a source of reference data. Since our focus is on the legs, one leg has been equipped with optical markers, providing reference position and orientation of the foot sensor, lower leg sensor, upper leg sensor and – not visible in the figure – the pelvis sensor.

Inertial data has been collected at 30 Hz. The sensors, however, run the strapdown integration algorithm discussed in Section 4 internally at 600 Hz to capture the high bandwidth of the measurement signals during impact, for



Fig. 3. Experimental setup where the human body is equipped with inertial sensors on different body segments. Optical markers for the reference system have been placed on the right foot sensor, right lower and upper leg sensors and – not visible in the figure – the pelvis.

instance during foot impact on the ground. To speed up the computations, the optimization algorithm itself has been run at a frequency of 10 Hz.

The optimization problem typically converges in a few iterations. To solve the problem for an experiment of 10 seconds takes about 5 minutes on an AMD X4 2.8 GHz processor for a first inefficient Matlab implementation of the algorithm. Initial tests with a C-implementation, however, show that speed improvements of up to 500 times are easily obtained. Taking into account that at the moment we postprocess the whole data set while for a real-time application a moving horizon can be used, we think that a real-time implementation of the algorithm is indeed quite possible.

The collected inertial data has been postprocessed used in the optimization problem (2) for a lower body configuration consisting of a set  $\mathcal{S}$  of 7 sensors placed on 7 body segments  $\mathcal{B}$ : both feet, both lower legs, both upper legs and the pelvis. The position of each sensor  $S_i \in \mathcal{S}$  on the body segment  $r_{S_i}^{B_{S_i}}$  has been manually measured. The orientations of the sensors on the body segments  $q^{B_{S_i}S_i}$  for all  $S_i \in \mathcal{S}$  have been determined by standing still in a pre-determined pose as described by Roetenberg et al. [2013]. The 7 body segments are connected by 6 joints  $\mathcal{J}$  of which the two knee joints are assumed to be hinge joints. Calibrating for the distances  $r_k^{B_j}$  amounts to defining the distances between the different joint centers which is again done by manual measuring. We acknowledge that this is an inaccurate calibration method and as future work we therefore plan to extend the algorithm to automatically estimate these calibration parameters.

Fig. 4 visualizes the pose of the lower body of a walking subject estimated using Algorithm 1 for parts of an experiment. Note that our experimental setup does allow for accurate *absolute* position estimates. The location of the different steps has therefore been corrected for one joint location using the position estimates from the optical reference system. The steps are taken from a short experiment and the optimization is run at 30 Hz for plotting purposes.



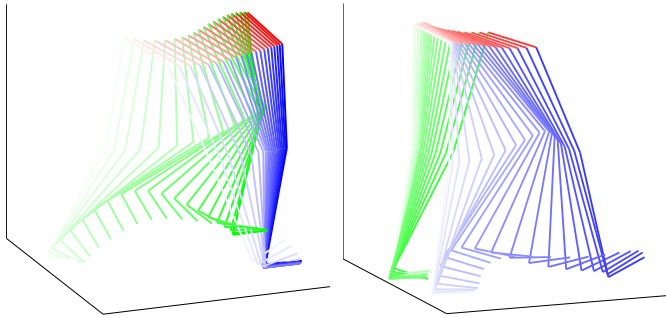


Fig. 4. Estimated pose of the lower body at different times during a step of the left leg (left) and the right leg (right). The view is chosen such that we view the subject from the right, and the right leg is depicted in blue, the left leg in green and the connection between the hips in red.

To compare our relative orientation results to those of the optical reference system, we focus on the estimated joint angle of the right knee during an experiment of around 37 seconds. Joint angles are defined as the angle between two connected body segments at the joint center. For the knee joint, the bending of the knee during walking is referred to as flexion/extension. The rotation around the other two axes (abduction/adduction and internal/external rotation) are generally quite small for this joint. Because it is not possible to observe the joint center and sensors/markers are generally placed on the soft tissue instead of on the bone, computation of joint angles depends on a model of the joint locations in the body. Theoretically, it is possible to estimate the joint angle from the orientation results of the sensor if the exact location of the sensors with respect to the joints is known, i.e. in case of a perfect calibration, and if the sensors would be rigidly attached to the bone. In practice this is clearly not possible. However, since both the inertial sensors and the optical reference markers are placed on the same location on the body segments as shown in Fig. 3, it is still possible to compare the angles to assess the quality of our estimates.

To be able to compare our joint angle estimates to those of the reference system, a coordinate frame alignment between the sensor coordinate frame and the coordinate frame of the optical markers needs to be performed. This has been done as described by Hol [2011]. Note that due to limited excitation of the upper leg sensor, it was not possible to do this alignment based on the sensor signals. Instead, the alignment has been performed based on the joint angle estimates. The joint angle estimates from our algorithm can be seen to match the joint angles from the optical reference system. A more quantitative analysis can be performed when the calibration parameters are properly estimated and position aiding is included. Note that due to the limited size of the measurement volume of the optical reference system, the movements are quite restricted and at some time instances in the experiment the optical reference data is not present.

From our optimization algorithm it is also possible to estimate the joint angles from the angles of the body segments. These are included in red in Fig. 5. There is

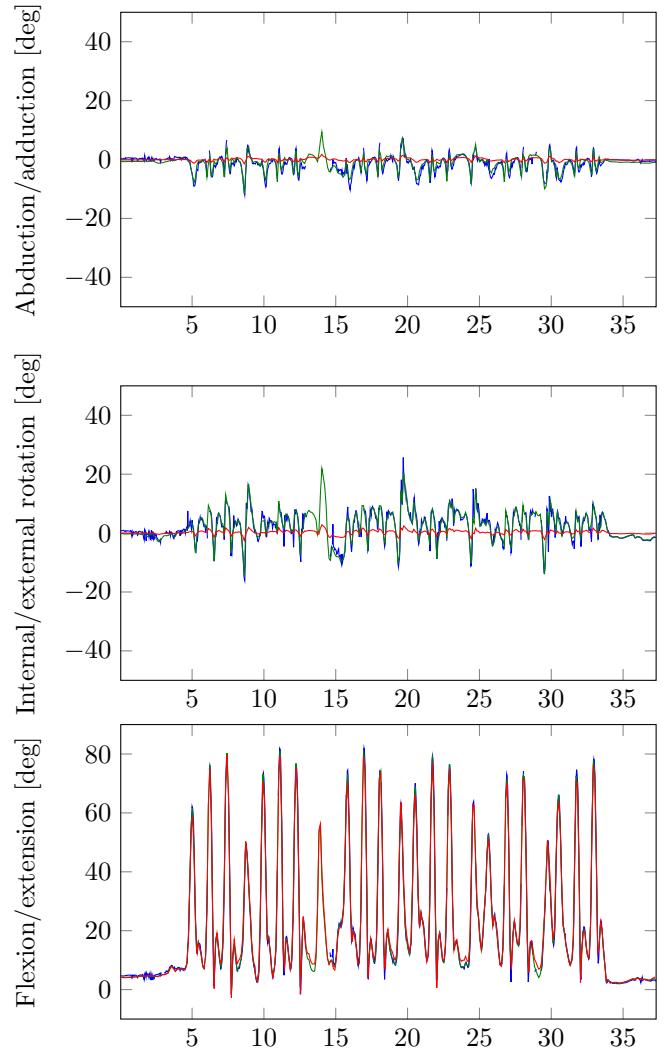


Fig. 5. Knee joint angles for the right knee for an experiment consisting of 23 steps. The optical reference data is plotted in blue, the joint angle estimated from the sensor's orientations, using our algorithm is plotted in green, the joint angle from the body segment orientations is plotted in red. Best viewed in color.

no validation for the angles obtained in this way, but the estimated abduction/adduction and internal/external rotation are considerably closer to zero, as we would expect from our knowledge that these rotations are quite small.

## 7. CONCLUSIONS AND FUTURE WORK

An optimization approach to inertial human body motion capture has been developed, capable of estimating the relative position and orientation of the body segments. Experimental results show that the algorithm works well, quickly converging to a feasible solution and resulting in drift-free joint angle estimates which match the joint angles from an optical reference system.

We plan to extend the approach to also estimate the calibration parameters and to include position aiding in the form of zero velocity updates at stand still and ultra-wideband position aiding [Hol et al., 2009]. This will also allow a more quantitative analysis of the results.

Future work also includes adding more body segments, modeling of non-Gaussian noise where appropriate and implementing a moving horizon estimation version of the algorithm.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr. Daniel Roetenberg for valuable discussions and Raymond Zandbergen for implementation support.

## REFERENCES

- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- J. Callmer. *Autonomous Localization in Unknown Environments*. PhD thesis, Linköping University, Sweden, June 2013.
- G. Cooper, I. Sheret, L. McMillan, K. Siliverdis, N. Sha, D. Hodgins, L. Kenney, and D. Howard. Inertial sensor-based knee flexion/extension angle estimation. *Journal of Biomechanics*, 42(16):2678 – 2685, 2009.
- J.L. Crassidis, F. Landis Markley, and Y. Cheng. A survey of nonlinear attitude estimation methods. *Journal of Guidance, Control, and Dynamics*, 30(1):12–28, 2007.
- N. El-Sheimy, H. Hou, and X. Niu. Analysis and modeling of inertial sensors using Allan variance. *IEEE Transactions on Instrumentation and Measurement*, 57(1):140–149, 2008.
- J. Favre, B.M. Jolles, R. Aissaoui, and K. Aminian. Ambulatory measurement of 3D knee joint angle. *Journal of biomechanics*, 41(5):1029–1035, 2008.
- G. Grisetti, R. Kummerle, C. Stachniss, U. Frese, and C. Hertzberg. Hierarchical optimization on manifolds for online 2D and 3D mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 273–278, Anchorage, Alaska, 2010.
- F. Gustafsson. *Statistical Sensor Fusion*. Studentlitteratur, 2012.
- J.D. Hol. *Sensor Fusion and Calibration of Inertial Sensors, Vision, Ultra-Wideband and GPS*. PhD thesis, Linköping University, Sweden, June 2011.
- J.D. Hol, F. Dijkstra, H. Luinge, and T.B. Schön. Tightly coupled UWB/IMU pose estimation. In *Proceedings of the IEEE International Conference on Ultra-Wideband (ICUWB)*, pages 688–692, Vancouver, Canada, September 2009.
- D.H. Kang, Y.J. Jung, A.J. Park, and J.W. Kim. Human body motion capture system using magnetic and inertial sensor modules. In *Proceedings of the 5th international universal communication symposium (IUCS)*, Gumi, Korea, October 2011.
- H.J. Luinge. *Inertial Sensing of Human Movement*. PhD thesis, University of Twente, Enschede, the Netherlands, October 2002.
- H.J. Luinge, P.H. Veltink, and C.T.M. Baten. Ambulatory measurement of arm orientation. *Journal of Biomechanics*, 40(1):78–85, 2007.
- J. Mattingley and S. Boyd. Real-time convex optimization in signal processing. *IEEE Signal Processing Magazine*, 27(3):50–61, 2010.
- T.B. Moeslund, A. Hilton, and V. Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2-3):90 – 126, 2006.
- C.V. Rao, J.B. Rawlings, and J.H. Lee. Constrained linear state estimation - a moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.
- D. Roetenberg, H.J. Luinge, and P. Slycke. Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors, May 2013.
- P.G. Savage. Strapdown inertial navigation integration algorithm design part 1: Attitude algorithms. *Journal of Guidance, Control and Dynamics*, 21(1):19–28, 1998a.
- P.G. Savage. Strapdown inertial navigation integration algorithm design part 2: Velocity and position algorithms. *Journal of Guidance, Control and Dynamics*, 21(2):208–221, 1998b.
- G. Welch and E. Foxlin. Motion tracking: No silver bullet, but a respectable arsenal. *IEEE Computer Graphics and Applications*, 22(6):24–38, 2002.
- O.J. Woodman. *Pedestrian localisation for indoor environments*. PhD thesis, University of Cambridge, United Kingdom, September 2010.
- Xsens Technologies B.V. <http://www.xsens.com>, Accessed on November 17, 2013.
- Q. Yuan and I.-M. Chen. 3-D localization of human based on an inertial capture system. *IEEE Transactions on Robotics*, 29(3):806–812, 2013.
- X. Yun and E.R. Bachmann. Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking. *IEEE Transactions on Robotics*, 22(6):1216–1227, 2006.