Proceedings of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

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

$$\begin{aligned} \max_{z} \quad & p(z \mid y_{1:N}) \\ \text{s.t.} \quad & c_{\mathrm{e}}(z) = 0, \end{aligned} \tag{1}$$

where $c_{\mathrm{e}}(z)$ represents the equality constraints. In our problem, z consists of both static parameters θ 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

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

Obtaining the MAP estimate thus amounts to solving a constrained optimization problem where the constraints $c_{\mathrm{bio}}(z)$ 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 L aligned with the local gravity vector, with the z-axis pointing up. The