Full Body Human Motion Estimation on Lie Groups Using 3D Marker Position Measurements

Josip Ćesić*, Vladimir Joukov‡, Ivan Petrović* and Dana Kulić‡

***Abstract*— This paper proposes a new algorithm for full body human motion estimation using 3D marker position measurements. The joints are represented with Lie group members, including special orthogonal groups $\mathsf{SO}(2)$ and $\mathsf{SO}(3)$, and a special euclidean group $\mathsf{SE}(3)$. We employ the Lie Group Extended Kalman Filter (LG-EKF) for stochastic inference on groups, thus explicitly accounting for the non-euclidean geometry of the state space, and provide the derivation of the LG-EKF recursion for articulated motion estimation. We evaluate the performance of the proposed algorithm in both simulation and on real-world motion capture data, comparing it with the Euler angles based EKF. The results show that the proposed filter significantly outperforms the Euler angles based EKF, since it estimates human motion more accurately and is not affected by gimbal lock.**

I. Introduction

Human bodies have evolved to perform complex manipulation and locomotion tasks. We are able to accomplish very intricate movements, carry light and heavy loads, achieve energy efficient locomotion at various speeds, reject disturbances, and adapt to environment constraints. Inspired by the human body, robotics researchers aim to develop systems with similar capabilities. To design a humanoid that can perform as well as a person, researchers must first capture and analyze human motion. Accurate pose estimation allows the design of controllers to simulate human like movements on a robot through motion re-targeting and imitation learning. In human-robot interaction the participant's pose must be known to guarantee safety and to allow collaborative tasks. Finally, to improve the performance of assistive devices in rehabilitation or to enhance user's capabilities with an exoskeleton, the system must be able to reproduce human like movements [1].

Optical motion capture is a method to record the movements from body worn markers observed by multiple cameras. The 3D positions of the markers are extracted from the images using the relative positions of the cameras to each other and are analyzed to compute the pose.

Typically, a kinematic model of the participant is defined based on anthropomorphic tables or by measurement and markers are assumed rigidly attached to the skeleton links. Unfortunately, for a full body skeletal model, there is no closed form solution for the inverse kinematics (IK). Differentiating the positions of the attached markers with respect to the joint angles and forming a Jacobian matrix allows to iteratively solve for joint angles using the pseudoinverse of the Jacobian. In singular configurations the Jacobian is not invertible. It is possible to include a non-zero damping constant in the least squares minimization to maintain full rank; various damping factors have been proposed [2].

The Jacobian inverse based methods do not account for stochastic error in marker position measurements, are greatly affected by outlier measurements, and are not capable of predicting future poses. By treating the skeleton pose as a state and 3D marker positions as measurements, recursive stochastic estimators can be used to help reduce the effect of stochastic marker position errors. Including the joint positions, velocities, and accelerations in the motion model of the filter helps to maintain correct pose estimate during short term occlusions. Various stochastic filters have been proposed for IK, such as the Smart Sampling Kalman Filter [3] and the Unscented Kalman Filter [4]. The filtering approach can even be used to perform estimation from unlabeled markers [5]. Bonnet *et al.* modelled not only kinematics but also the dynamics of a human body within an EKF to estimate the pose and dynamic parameters [6].

In the aforementioned methods the kinematic models are rigid links connected with joints that may be rotational, translational, or spherical. All of these formulations are representations of transformations in the euclidean space. However, human motion and many other types of motion of interest in robotics do not occur in Euclidean space, but rather arise on curved geometries often called manifolds. By using the manifold representations, the overall performance of wide variety of applications can be significantly improved [7]–[9]. In particular, the attitude of an object can be modelled as a special orthogonal group $\mathsf{SO}(n)$, $n = 2, 3$, while the pose can be modelled as a special euclidean group $\mathsf{SE}(n)$, $n = 2, 3$ [7]. Notably, both $\mathsf{SO}(n)$ and $\mathsf{SE}(n)$ belong to a family of matrix Lie groups. Recently, several theoretically rigorous approaches for filtering on manifolds have been proposed. In [10] the authors proposed an EKF able to perform estimation respecting the geometry of matrix Lie groups. Alongside, the unscented transform-based [11] and the particle-based [12] approaches have also attracted significant attention.

The benefit of manifolds for human action recognition has already been explored in the literature. In [13] the

This work has been supported from the Unity Through Knowledge Fund under the project *Cooperative Cloud based Simultaneous Localization and Mapping in Dynamic Environments* (cloudSLAM) and by the Ministry of Science, Education and Sports of the Republic of Croatia under the grant *Center of Research Excellence for Data Science and Cooperative Systems* (CoE ACROSS).

* Josip Ćesić and Ivan Petrović are with the University of Zagreb, Faculty of Electrical Engineering and Computing, Laboratory for Autonomous Systems and Mobile Robotics, Croatia. {josip.cesic@fer.hr, ivan.petrovic@fer.hr}

‡ Vladimir Joukov and Dana Kulić are with the University of Waterloo, Department of Electrical and Computer Engineering, Adaptive Systems Laboratory, Canada. {vjoukov@uwaterloo.ca, dkulic@uwaterloo.ca}

authors exploited the manifold structure by relying on the particle filter for learning purposes, while in [14] the authors use different manifolds as priors for manifold learning. Devanne *et al.* have used a spatio-temporal modeling of trajectories in a Riemannian manifold for action recognition purposes [15]. Recently, Brossette *et al.* have proposed the posture generation problem that encompasses non-Euclidean manifolds as well [16].

In this paper, we propose an algorithm for human motion estimation on Lie groups, which uses 3D marker position measurements. We explicitly account for the geometry of the state space and apply Lie group EKF (LG-EKF) for stochastic inference on Lie groups. We employ a constant acceleration model [17] in the motion prediction step and derive the update and observation equations for positional measurements. We compare the performance of the proposed approach with the Euler angles-based EKF, and show that the proposed algorithm achieves significantly better performance in both simulations and real-world experiments.

The paper is organized as follows. In Sec. II we present the theoretical preliminaries addressing the association of uncertainties to Lie groups, and provide the basic relations needed for forward kinematics of articulated bodies with groups. In Sec. III we derive the proposed estimation approach. In Sec. IV we describe the Euler angle-based approach, while in Sec. V we present the validation results.

II. Mathematical background

In this section we provide the mathematical background for performing human motion estimation on matrix Lie groups. We first discuss a human body modeling approach and the corresponding state space construction, and after provide the relations for manipulating the required Lie group members.

A. Construction of the state space

Before proceeding to filtering, we first construct the state space for representing a human that models body flexibility to a satisfactory level. Therefore we determine the appropriate Lie Group representation for each joint based on its mobility. For example, 1 DoF revolute joints are represented with a special orthogonal group $\mathsf{SO}(2)$, while 3 DoF spherical joints are modelled with a special orthogonal group $\mathsf{SO}(3)$. To localize the human in 6 DoF space, we use a special euclidean group member $\mathsf{SE}(3)$ for connecting the origin of the space with the base of the body, modeling both translational and rotational motion. Finally, the state of the system modelling a human is constructed by concatenating Lie group members via a Cartesian product, starting with $\mathsf{SE}(3)$, and extending with either $\mathsf{SO}(2)$ or $\mathsf{SO}(3)$ groups.

For example, a human leg can be constructed as

$$\mathsf{SE}(3) \times \mathsf{SO}(3) \times \mathsf{SO}(2) \times \mathsf{SO}(2) \times \mathsf{SO}(2)\,. \tag{1}$$

Here, the first term represents the 3D position and orientation of the waist with respect to the reference frame, the second term represents the hip as a spherical joint, the third describes the knee, while the last two represent the two dimensional ankle as shown in Fig. 1 (left).

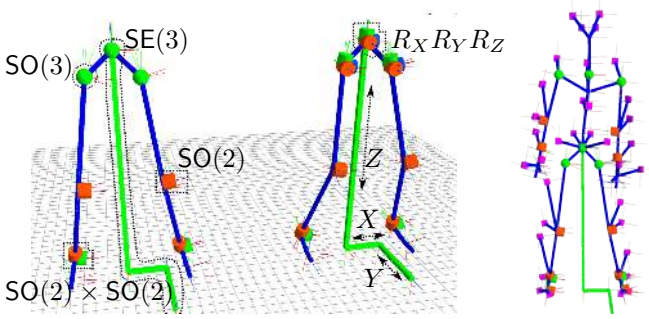

Fig. 1: Left: Lower body kinematic model joints represented by their respective group members. Middle: Same lower body in prismatic and revolute (Euler angle) joint representation. Right: Full body Lie Group model with attached markers.

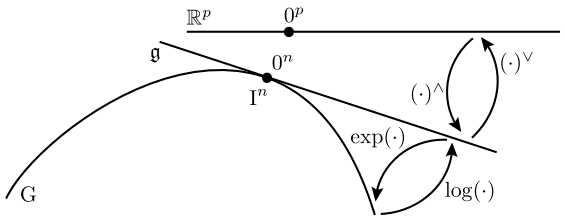

Fig. 2: An illustration of mappings within the triplet of Lie group G, Lie algebra $\mathfrak{g}$ and the Euclidean space $\mathbb{R}^p$.

B. Lie groups and Lie algebra

We now introduce the concept of Lie groups and Lie algebra as prerequisites for estimation on Lie groups [18].

Generally, a Lie group G is a group which has the structure of a smooth manifold. Group operators, composition and inversion, are smooth operations, given simply as matrix multiplication and inversion. Lie algebra $\mathfrak{g}$ elements represent a tangent space of a group at the identity element [19]. In particular, a Lie algebra is an open neighborhood around $\mathbf{0}^p$ in the tangent space of G at the identity I^n. The matrix exponential $\exp_{\mathsf{G}}$ and matrix logarithm $\log_{\mathsf{G}}$ establish a local diffeomorphism between G and $\mathfrak{g}$ as

$$\exp_{\mathsf{G}} : \mathfrak{g} \to \mathsf{G} \quad \text{and} \quad \log_{\mathsf{G}} : \mathsf{G} \to \mathfrak{g}. \tag{2}$$

The Lie algebra $\mathfrak{g}$ associated to a p-dimensional matrix Lie group $\mathsf{G} \subset \mathbb{R}^{n\times n}$ is a p-dimensional vector space defined by a basis consisting of p real matrices E_r, $r = 1, .., p$, often referred to as generators [20]. A linear isomorphism between $\mathfrak{g}$ and $\mathbb{R}^p$ is given by

$$[\cdot]^\vee_{\mathsf{G}} : \mathfrak{g} \to \mathbb{R}^p \quad \text{and} \quad [\cdot]^\wedge_{\mathsf{G}} : \mathbb{R}^p \to \mathfrak{g}. \tag{3}$$

An illustration of the above mappings is given in Fig. 2.

In addition, in Lie group based calculus we need two more operators – adjoint representation of a Lie group, denoted as Ad_{G} and Lie algebra ad_{G}. More detailed discussion on adjoints and the used notation can be found in [18] and [10], respectively.

C. Concentrated Gaussian distribution

To make use of EKF on Lie groups, the Gaussian error distribution covariance must be established. Distribution on the group tightly focused around the identity element X^I

can be expressed on the Lie algebra [21] with probability density function given as

$$p(X^I) = \beta \exp\left(-\frac{1}{2}[\log_{\mathsf{G}}(X^I)]_{\mathsf{G}}^{\vee^T} P^{-1} [\log_{\mathsf{G}}(X^I)]_{\mathsf{G}}^{\vee}\right),$$

where β is a normalizing constant and P is a positive definite matrix. If $\epsilon \triangleq [\log_{\mathsf{G}}(X^I)]_{\mathsf{G}}^{\vee}$ is tightly focused, it can be described with a classical Gaussian $\epsilon \sim \mathcal{N}_{\mathbb{R}^p}(\mathbf{0}^{p\times 1}, P)$. The distribution of random variable X^I can be translated over G by using the left action of the Lie group, and finally a random variable X can be seen as

$$X = \mu \exp_{\mathsf{G}}(\epsilon_{\mathsf{G}}^{\wedge}) \text{ , with } X \sim \mathcal{G}(\mu, P), \tag{4}$$

where $\mathcal{G}$ denotes the so called concentrated Gaussian distribution (CGD) [21]. For a more formal introduction, the interested reader is referred to [18].

D. Special orthogonal group $\mathsf{SO}(2)$

The $\mathsf{SO}(2)$ group represents a rotation around a single axis:

$$\mathsf{SO}(2) = \left\{X \subset \mathbb{R}^{2\times 2} \,|\, X^{\mathrm{T}}X = I,\, \det(X) = 1\right\}. \tag{5}$$

For a euclidean space vector consisting of an angle $x = \phi$, the Lie algebra $\mathfrak{so}(2)$, is given as

$$x_{\mathsf{SO}(2)}^{\wedge} = \begin{bmatrix} 0 & -\phi \\ \phi & 0 \end{bmatrix} \in \mathfrak{so}(2). \tag{6}$$

where $(\cdot)_{\mathsf{SO}(2)}^{\wedge} : \mathbb{R}^1 \to \mathfrak{so}(2)$. Its inverse, $(\cdot)_{\mathsf{SO}(2)}^{\vee} : \mathfrak{so}(2) \to \mathbb{R}^1$, follows trivially from relation (6). The exponential for $\mathsf{SO}(2)$, performing $\exp_{\mathsf{SO}(2)} : \mathfrak{so}(2) \to \mathsf{SO}(2)$, is given as

$$\exp_{\mathsf{SO}(2)}(x_{\mathsf{SO}(2)}^{\wedge}) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \tag{7}$$

while the inverse operator, $\log_{\mathsf{SO}(2)} : \mathsf{SO}(2) \to \mathfrak{so}(2)$, can be evaluated from (7). Due to the commutativity of $\mathsf{SO}(2)$, the adjoint operators are given as

$$\mathrm{Ad}_{\mathsf{SO}(2)}(X) = 1 \text{ and } \mathrm{ad}_{\mathsf{SO}(2)}(x) = 0. \tag{8}$$

These properties will greatly simplify the LG-EKF formulae.

E. Special orthogonal group $\mathsf{SO}(3)$

The $\mathsf{SO}(3)$ group represents an orientation of a rigid body in 3D space, and is defined as

$$\mathsf{SO}(3) = \left\{X \subset \mathbb{R}^{3\times 3} \,|\, X^{\mathrm{T}}X = I,\, \det(X) = 1\right\}. \tag{9}$$

For a euclidean joint space vector $x = [\phi_1\ \phi_2\ \phi_3]^{\mathrm{T}}$, the Lie algebra $\mathfrak{so}(3)$ is given as a skew symmetric matrix

$$x_{\mathsf{SO}(3)}^{\wedge} = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix} \in \mathfrak{so}(3). \tag{10}$$

where $(\cdot)_{\mathsf{SO}(3)}^{\wedge} : \mathbb{R}^3 \to \mathfrak{so}(3)$. Its inverse, $(\cdot)_{\mathsf{SO}(3)}^{\vee} : \mathfrak{so}(3) \to \mathbb{R}^3$, follows trivially from (10). The exponential for $\mathsf{SO}(3)$, performing mapping $\exp_{\mathsf{SO}(3)} : \mathfrak{so}(3) \to \mathsf{SO}(3)$, is given as

$$\begin{aligned}\exp_{\mathsf{SO}(3)}(x_{\mathsf{SO}(3)}^{\wedge}) = &\cos(|x|)I^3 + \\ &+ (1-\cos(|x|))\frac{xx^{\mathrm{T}}}{|x|^2} + \sin(|x|)\frac{x_{\mathsf{SO}(3)}^{\wedge}}{|x|}.\end{aligned} \tag{11}$$

The logarithm, performing mapping $\log_{\mathsf{SO}(3)} : \mathsf{SO}(3) \to \mathfrak{so}(3)$, is given as

$$\begin{aligned}\log_{\mathsf{SO}(3)}(X) &= \frac{\theta}{2\sin(\theta)}(X - X^{\mathrm{T}}) \\ \text{s.t.} \quad & 1 + 2\cos(\theta) = \mathrm{Tr(X)} \\ & \begin{cases} \theta \neq 0 & -\pi < \theta < \pi \\ \theta = 0 & \log(X) = 0 \end{cases}.\end{aligned} \tag{12}$$

The adjoints $\mathrm{Ad}_{\mathsf{SO}(3)}$ and $\mathrm{ad}_{\mathsf{SO}(3)}$ are respectively given as

$$\mathrm{Ad}_{\mathsf{SO}(3)}(X) = X \text{ and } \mathrm{ad}_{\mathsf{SO}(3)}(x) = x_{\mathsf{SO}(3)}^{\wedge}. \tag{13}$$

F. Special euclidean group $\mathsf{SE}(3)$

The group $\mathsf{SE}(3)$ describes 6 DoF rigid body pose and is formed as a semi-direct product of the euclidean space vector $\mathbb{R}^3$ and the special orthogonal group $\mathsf{SO}(3)$[1], corresponding to translational and rotational parts, respectively. This group is defined as

$$\mathsf{SE}(3) = \left\{ \begin{pmatrix} R & t \\ \mathbf{0} & 1 \end{pmatrix} \subset \mathbb{R}^{4\times 4} \,|\, \{R, t\} \in \mathsf{SO}(3) \times \mathbb{R}^3 \right\}.$$

For a euclidean space vector representing the pose of a rigid body consisting of a 3DoF position vector t and a 3DoF orientation vector ϕ, where $x = [t\ \phi]^{\mathrm{T}}$, the Lie algebra $\mathfrak{se}(3)$ is

$$x_{\mathsf{SE}(3)}^{\wedge} = \begin{bmatrix} \phi_{\mathsf{SO}(3)}^{\wedge} & t \\ \mathbf{0} & 0 \end{bmatrix} \in \mathfrak{se}(3). \tag{14}$$

where $(\cdot)_{\mathsf{SE}(3)}^{\wedge} : \mathbb{R}^6 \to \mathfrak{se}(3)$. Its inverse, $(\cdot)_{\mathsf{SE}(3)}^{\vee} : \mathfrak{se}(3) \to \mathbb{R}^6$, follows trivially from (14). The exponential for $\mathsf{SE}(3)$, performing mapping $\exp_{\mathsf{SE}(3)} : \mathfrak{se}(3) \to \mathsf{SE}(3)$, is given as

$$\exp_{\mathsf{SE}(3)}(x_{\mathsf{SE}(3)}^{\wedge}) = \begin{bmatrix} C & Lt \\ \mathbf{0} & 1 \end{bmatrix} \tag{15}$$

$$C = \exp_{\mathsf{SO}(3)}(\phi_{\mathsf{SO}(3)}^{\wedge})$$

$$L = \frac{\sin(|\phi|)}{|\phi|}I^3 + (1 - \frac{\sin(|\phi|)}{|\phi|})\frac{\phi\phi^{\mathrm{T}}}{|\phi|^2} + \frac{1-\cos(|\phi|)}{|\phi|^2}\phi_{\mathsf{SO}(3)}^{\wedge}.$$

The logarithm, performing mapping $\log_{\mathsf{SE}(3)} : \mathsf{SE}(3) \to \mathfrak{se}(3)$, is calculated by deconstructing X, and determining ϕ by using (12). Then, from (15) we can determine t.

In order to determine the adjoints for $\mathsf{SE}(3)$, we need to deconstruct the state $X \in \mathsf{SE}(3)$ and vector $x \in \mathbb{R}^6$. Firstly, we extract the rotation part C and translation part t from X, and secondly, we split the translation part t and orientation part ϕ from x. Then, the adjoints $\mathrm{Ad}_{\mathsf{SE}(3)}$ and $\mathrm{ad}_{\mathsf{SE}(3)}$ are

$$\mathrm{Ad}_{\mathsf{SE}(3)}(X) = \begin{bmatrix} C & tC \\ \mathbf{0} & C \end{bmatrix}, \ \mathrm{ad}_{\mathsf{SE}(3)}(x) = \begin{bmatrix} \phi_{\mathsf{SO}(3)}^{\wedge} & t_{\mathsf{SO}(3)}^{\wedge} \\ \mathbf{0} & \phi_{\mathsf{SO}(3)}^{\wedge} \end{bmatrix}.$$

We next present the new human motion estimation method based on the LG-EKF.

[1]The euclidean space can be formed only by employing direct product, while other ways to concatenate Lie groups also exist, i.e., semi-direct product, twisted product, etc.

III. Human motion estimation on Lie groups

The LG-EKF performs motion prediction and measurement update steps recursively, assuming a constant acceleration model (CA) [17] for each joint.

A. Motion prediction step

The LG-EKF approach assumes the motion model of the system can be described with the following equation

$$X_{k+1} = f(X_k, n_k) = X_k \exp_{\mathsf{G}}\left([\hat{\Omega}_k + n_k]^{\wedge}_{\mathsf{G}}\right), \quad (16)$$

where $X_k \in \mathsf{G}$ is the state of the system at time k, G is a p-dimensional Lie group, $n_k \sim \mathcal{N}_{\mathbb{R}^p}(\mathbf{0}^{p\times 1}, Q_k)$ is zero mean white Gaussian noise with covariance Q_k and $\hat{\Omega}_k = \Omega(X_k) : \mathsf{G} \to \mathbb{R}^p$ is a non-linear $\mathcal{C}^2$ function.

For example, assuming a CA motion model and considering a single $\mathsf{SO}(2)$ joint with associated angular velocity and angular acceleration, the state would be given by $X \in \mathsf{G} = \mathsf{SO}(2) \times \mathbb{R}^1 \times \mathbb{R}^1$, and

$$\hat{\Omega}_k = \begin{bmatrix} T\dot{q}_k + \frac{T^2}{2}\ddot{q}_k \\ T\ddot{q}_k \\ 0 \end{bmatrix} \in \mathbb{R}^3, \; n_k = \begin{bmatrix} \frac{T^2}{2} n_k^a \\ T n_k^a \\ n_k^a \end{bmatrix} \in \mathbb{R}^3, \quad (17)$$

where q_k, $\dot{q}_k$ and $\ddot{q}_k$ are the angle, angular velocity and angular acceleration represented in tangential space, respectively[2]. The term n_k^a represents the acceleration increment during the k-th sampling period [17].

In general, the state of the system X is formed by using direct (Cartesian) product between the group members, i.e., by placing them block-diagonally. Then, after applying $\exp_{\mathsf{G}}$ or $\log_{\mathsf{G}}$, the element will stay in the block diagonal arrangement. The motion model $\hat{\Omega}_k$ can be seen as representing an addition to the current state, and for N joints it is given as $\hat{\Omega}_k = [\hat{\Omega}_k^1 \; \hat{\Omega}_k^2 \; ... \; \hat{\Omega}_k^N]^{\mathsf{T}}$. The motion model and the process noise associated with the i-th joint, i.e., $\hat{\Omega}_k^i$ and n_k^i, are elements of euclidean space $\mathbb{R}^r$, where $r = 3 \times$ (# DoF) since position, velocity and acceleration are included. Hence, for the associated group member $\mathsf{SO}(2)$, $\mathsf{SO}(3)$ and $\mathsf{SE}(3)$, the coefficient is $r = 3$, $r = 9$ and $r = 18$.

We assume the posterior distribution at step $k-1$ follows the concentrated Gaussian distribution assumption $\mathcal{G}(\mu_{k-1}, P_{k-1})$. The mean propagation of the LG-EKF is then governed by

$$\mu_{k+1|k} = \mu_k \exp_{\mathsf{G}}\left([\hat{\Omega}_k]^{\wedge}_{\mathsf{G}}\right), \quad (18)$$

while the covariance prediction is governed by

$$P_{k+1|k} = \mathcal{F}_k P_k \mathcal{F}_k^{\mathsf{T}} + \Phi_{\mathsf{G}}(\hat{\Omega}_k) Q_k \Phi_{\mathsf{G}}(\hat{\Omega}_k)^{\mathsf{T}}. \quad (19)$$

The operator $\mathcal{F}_k$ represents a matrix Lie group equivalent to the Jacobian of $f(X_k, n_k)$, and is calculated by

$$\begin{aligned} \mathcal{F}_k &= \mathrm{Ad}_{\mathsf{G}}\left(\exp_{\mathsf{G}}\left([-\hat{\Omega}_k]^{\wedge}_{\mathsf{G}}\right)\right) + \Phi_{\mathsf{G}}(\hat{\Omega}_k)\mathscr{C}_k \\ \mathscr{C}_k &= \frac{\partial}{\partial \epsilon}\Omega\left(\mu_k \exp_{\mathsf{G}}(\epsilon^{\wedge}_{\mathsf{G}})\right)_{|\epsilon=0}. \end{aligned} \quad (20)$$

[2]Euclidean space $\mathbb{R}^p$ belongs to a family of Lie groups, while for constructing G we employ its matrix representation obtained by matrix embedding. It is also a subgroup of $\mathsf{SE}(n)$ where a pure translation is employed [18].

$\mathscr{C}_k$ represents the linearisation term where the argument of the motion model is the current state X_k with an incremental perturbation additively added in each of the p directions. Contrary to the conventional EKF, a linear additive process noise injects the system as a function of the current state of the system over the transformation $\Phi_{\mathsf{G}}(\hat{\Omega}_k) Q_k \Phi_{\mathsf{G}}(\hat{\Omega}_k)^{\mathsf{T}}$, where Φ_{G} appears due to the displacement of the tangential space during the prediction step, and is given as

$$\Phi_{\mathsf{G}}(v) = \sum_{i=0}^{\infty} \frac{(-1)^i}{(i+1)!} \mathrm{ad}_{\mathsf{G}}(v)^i, \; v \in \mathbb{R}^p. \quad (21)$$

B. Measurement update step

We next derive the update step by employing position measurements of markers attached to a human body obtained by a motion capture system. The markers are assumed to be rigidly attached to a predetermined skeletal model. The discrete measurement model on the matrix Lie group is modelled as

$$Z_{k+1} = h(X_{k+1}) \exp_{\mathsf{G}'}\left([m_{k+1}]^{\wedge}_{\mathsf{G}'}\right), \quad (22)$$

where $Z_{k+1} \in \mathsf{G}'$, $h : \mathsf{G} \to \mathsf{G}'$ is a $\mathcal{C}^1$ function and $m_{k+1} \sim \mathcal{N}_{\mathbb{R}^q}(\mathbf{0}^{q\times 1}, R_{k+1})$ is zero-mean white Gaussian noise with covariance R_{k+1}. The measurement function, in our marker based approach, is given as $h(X_{k+1}) = \mathrm{diag}\{h(X_{k+1})^1, h(X_{k+1})^2, .., h(X_{k+1})^M\}$, where M block-diagonally placed measurement components correspond to M marker position measurements, and hence the measurement space is given as $\mathsf{G}' = \mathbb{R}^{3M}$.

The update step of the filter strongly resembles the standard EKF update procedure, relying on the Kalman gain K_{k+1} and innovation vector ν_{k+1} calculated as

$$\begin{aligned} K_{k+1} &= P_{k+1|k}\mathcal{H}_{k+1}^{\mathsf{T}}\left(\mathcal{H}_{k+1}P_{k+1|k}\mathcal{H}_{k+1}^{\mathsf{T}} + R_{k+1}\right)^{-1} \\ \nu_{k+1} &= K_{k+1}\left(\left[\log_{\mathsf{G}'}\left(h(\mu_{k+1|k})^{-1} Z_{k+1}\right)\right]^{\vee}_{\mathsf{G}'}\right). \end{aligned} \quad (23)$$

The matrix $\mathcal{H}_{k+1}$ can be seen as a matrix Lie group equivalent to the Jacobian of $h(X_{k+1})$, and is given as

$$\mathcal{H}_{k+1} = \frac{\partial}{\partial \epsilon}\left[\log_{\mathsf{G}'}\left(h(\mu_{k+1|k})^{-1} h(\mu^{\epsilon}_{k+1|k})\right)\right]^{\vee}_{\mathsf{G}'\;|\epsilon=0},$$

where $h(\mu^{\epsilon}_{k+1|k}) = h(\mu_{k+1|k}\exp_{\mathsf{G}}(\epsilon^{\wedge}_{\mathsf{G}}))$, describes the variation of markers' positions for an infinitesimal motion ϵ. We now evaluate the part of $\mathcal{H}_{k+1}$ corresponding to the i-th marker's measurement Z^i_{k+1}. This relation is given as

$$\begin{aligned} \mathcal{H}^i_{k+1} =& \frac{\partial}{\partial \epsilon}\Big(\log_{\mathsf{G}'}\Big(h\left(\mathcal{K}^0_{s_i}(X_{k+1|k})\right)^{-1} \\ & h\left(\mathcal{K}^0_{s_i}(X^{\epsilon}_{k+1|k})\right)\Big)\Big)^{\vee}_{\mathsf{G}'\;|\epsilon=0} \\ =& \frac{\partial}{\partial \epsilon}\left(\log_{\mathsf{G}'}\left(\begin{bmatrix} I & \mathcal{K}^0_{s_i}(X^{\epsilon}_{k+1|k}) \begin{bmatrix} 0\\0\\0\\1 \end{bmatrix} \\ \mathbf{0} & \end{bmatrix}\right)\right)^{\vee}_{\mathsf{G}'\;|\epsilon=0}, \end{aligned} \quad (24)$$

where $\mathcal{K}^0_{s_i}(X_{k+1|k})$ stands for the forward kinematics of the i-th marker for a given predicted state $X_{k+1|k}$, while $\mathcal{K}^0_{s_i}(X^{\epsilon}_{k+1|k}) = \mathcal{K}^0_{s_i}\left(X_{k+1|k}\exp_{\mathsf{G}}(\epsilon^{\wedge}_{\mathsf{G}})\right)$ corresponds to the

forward kinematics for the infinitesimally perturbed state $X_{k+1|k}$. Note that the term $\mathcal{K}^0_{s_i}(X_{k+1|k})^{-1}$ vanishes after applying the partial derivatives over ϵ. We now decompose the kinematics term $\mathcal{K}^0_{s_i}(X_{k+1|k})$ into several parts as

$$\mathcal{K}^0_{s_i}(X_{k+1|k}) = \mathcal{K}^0_j(X_{k+1|k}) X^j_{k+1|k} \mathcal{K}^{j+1}_{s_i}(X_{k+1|k}), \quad (25)$$

where $\mathcal{K}^0_j(X_{k+1|k})$ represents the transformation from the base frame to joint j and $\mathcal{K}^{j+1}_{s_i}(X_{k+1|k})$ represents the transformation from joint $j+1$ towards sensor i.

Let us now consider a part of the $\mathcal{H}^i_{k+1}$ term relating the i-th measurement with the j-th joint, denoted as $\mathcal{H}^{i,j}_{k+1}$. Furthermore, let us assume the j-th joint is represented with an $\mathsf{SE}(3)$ term, hence covering the most general case, since $\mathsf{SO}(2)$ and $\mathsf{SO}(3)$ are simplifications of $\mathsf{SE}(3)$. Then, by exploiting results from [19], $\mathcal{H}^{i,j}_{k+1}$ can be expressed as

$$\begin{bmatrix} \mathcal{H}^{i,j,r}_{k+1} \\ 1 \end{bmatrix} = \mathcal{K}^0_j(X_{k+1|k}) X^j_{k+1|k} E^r \mathcal{K}^{j+1}_{s_i}(X_{k+1|k}) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

where E^r represents the r-th generator of $\mathsf{SE}(3)$ group, i.e., $r = 1, .., 6$ [21]. Each of the 6 generators represents an infinitesimal motion in one of the directions of $\mathsf{SE}(3)$ space, and $\mathcal{H}^{i,j}_{k+1} = \begin{bmatrix} \mathcal{H}^{i,j,1}_{k+1} \ ... \ \mathcal{H}^{i,j,6}_{k+1} \end{bmatrix}$. Since marker position measurements are only a function of the joint positions, the part of the $\mathcal{H}_{k+1}$ matrix relating measurements with velocity and acceleration components is filled with zero values.

Finally, the measurement update step is calculated as

$$\begin{aligned} \mu_{k+1} &= \mu_{k+1|k} \exp_{\mathsf{G}}\left([\nu_{k+1}]^{\wedge}_{\mathsf{G}}\right) \\ P_{k+1} &= \Phi_{\mathsf{G}}(\nu_{k+1}) \left(I^p - K_{k+1}\mathcal{H}_{k+1}\right) P_{k+1|k} \Phi_{\mathsf{G}}(\nu_{k+1})^{\mathrm{T}}. \end{aligned} \quad (26)$$

For a more formal derivation of the LG-EKF update, the interested reader is referred to [10].

IV. Euler angle based approach

The proposed approach is compared to conventional EKF applied to a standard kinematic model defined with revolute and prismatic joints [22]. Three perpendicular revolute joints (Euler angles) can be used to model human spherical joints such as the shoulder and the hip. The transformation between the world frame and the base of the body can be modelled with three prismatic and three perpendicular revolute joints, as shown in Fig. 1 (right). The state of the EKF is defined as the position q, velocity $\dot{q}$, and acceleration $\ddot{q}$ of the joints. Assuming constant acceleration the linear motion model is

$$\begin{aligned} q_{k+1} &= q_k + T\dot{q}_k + \frac{T^2}{2}\ddot{q}_k \\ \dot{q}_{k+1} &= \dot{q}_k + T\ddot{q}_k \\ \ddot{q}_{k+1} &= \ddot{q}_k \,. \end{aligned} \quad (27)$$

Treating the attached markers as end effectors, the measurement Jacobian for the i-th marker, H_i, is the velocity Jacobian in the base frame.

$$H^i = [Jv^{i1} Jv^{i2} \dots Jv^{in}] \quad (28)$$

$$Jv^{ij} = \begin{cases} z^j \times (o^i - o^j) & \text{for revolute joint } j \\ z^j & \text{for prismatic joint } j \end{cases} \quad (29)$$

where joint j is centered at o^j and actuates about z^j axis and o^i is the end effector position. With the Jacobians defined EKF can be set up to estimate the positions, velocities, and accelerations of all the joints in the kinematic model based on motion capture marker measurements.

V. Validation Results

We validate the proposed approach with three datasets. First, in simulation, we demonstrate the benefits of LG-EKF over EKF during highly dynamical movements whose motion is better described on the group and show that unlike EKF, LG-EKF is not affected by gimbal lock. Next, to show the benefits of $\mathsf{SO}(3)$ representation, we evaluate the performance of LG-EKF and EKF on real motion capture data of arm boxing movement. Finally, we perform full body estimation of a highly dynamic martial arts movement sequence to verify the effectiveness of the $\mathsf{SE}(3)$ joint connecting between world origin and the body base frame and demonstrate the overall benefits of LG-EKF over EKF.

A. Simulation Validation

1) Dynamic Motion: To test the convergence and estimation properties of LG-EKF, we simulate a human arm composed of the shoulder, elbow, and wrist joints, the state is an element of $\mathsf{SO}(3) \times \mathsf{SO}(2) \times \mathsf{SO}(3)$ group respectively. Two simulated motion capture markers are placed at the shoulder and elbow and 4 about the wrist. The kinematic chain is visualized in Fig. 3 (middle). We generate angular velocity on the group using a Fourier series with 5 harmonics and coefficients from a uni-variate distribution, the angular velocity is then propagated at 100 Hz according to the motion model defined in equation 16 with no additive noise. The simulated marker positions are computed with forward kinematics and Gaussian noise with standard deviation $\sigma_{\text{dev}} = 1$ mm is added to simulate errors in 3D marker measurement. This creates a highly dynamic motion as can be seen from the positions of the four wrist markers in Fig. 4. The measurement noise was set to 0.01 for both LG-EKF and EKF. No further tuning was performed to improve estimation of either filter, the initial covariances were set to identity and process noise for all states was 0.01.

To compare the estimate with the ground truth, we use the deviation from the identity matrix as the distance metric [23]

$$\mathcal{D}_F = \left\| I - R_e^{\mathrm{T}} R_{gt} \right\|_F \quad (30)$$

where R_e and R_{gt} are the estimated and ground truth rotation matrices of each joint and $\|\cdot\|_F$ denotes the Frobenius norm, which is funcionally equivalent to the geodesic on $\mathsf{SO}(3)$ [23]. Figure 5 shows the comparison in estimation of rotation matrices for each of the three joints

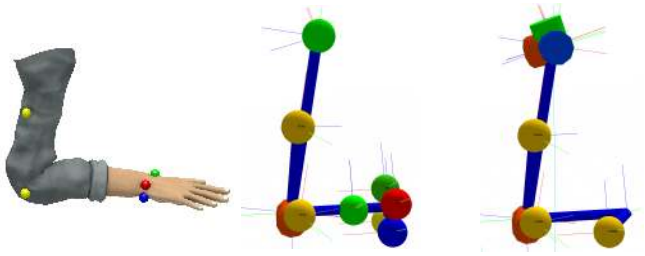

Fig. 3: Left: 3D Arm model showing simulation marker placement. Middle: Lie group-based arm model with attached markers for dynamic motion simulation. Right: Euler angle-based arm model for the CMU dataset (no wrist) with CMU markers attached.

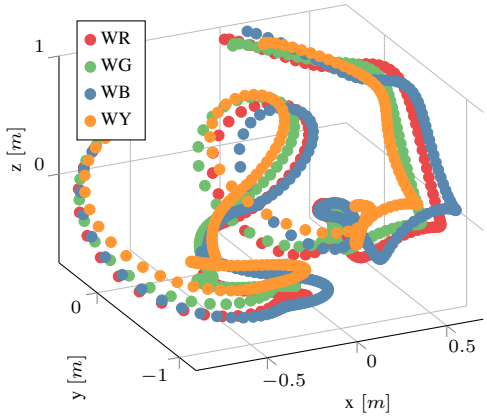

Fig. 4: Trajectories of wrist markers attached to the simulated arm model undergoing the generated highly dynamic motion over 2.5 s.

between LG-EKF and EKF using this distance metric. On average LG-EKF improves estimation over EKF by 20.9%. The observed improvement is composed of gimbal lock avoidance, described in the next section, and a better error covariance representation on the manifold.

2) Gimbal Lock: Any set of Euler angles will lose a degree of freedom when two of the rotation axes align [24], implying that in that configuration the rotation about the locked axis cannot be correctly estimated by EKF. Typically the order of the joint axes is carefully selected to try and avoid the lock, however in human motion estimation gimbal lock often takes place at the shoulder joint due to its high manoeuvrability. Unlike the Euler angle formulation, an $\mathsf{SO}(3)$ representation of the spherical joint does not suffer from gimbal lock and thus LG-EKF will accurately estimate any 3D rotation.

To demonstrate the benefits of LG-EKF over EKF during gimbal lock we simulate a single spherical joint at the origin with three motion capture markers attached at offsets of $[0.3, 0.1, 0]^{\mathrm{T}}$, $[0.3, -0.1, 0]^{\mathrm{T}}$, and $[0.3, 0, 0.1]^{\mathrm{T}}$ for full observability. To ensure continuation in position, velocity, and acceleration we use a quintic polynomial to generate a smooth trajectory, sampling at 200 Hz. First, the model experiences a 1 second rotation about the world y axis with initial position 0 rads and final positions $\frac{\pi}{2}$ rads and zero initial and final velocity and acceleration. Since the second joint of the Euler model is aligned with the y axis this

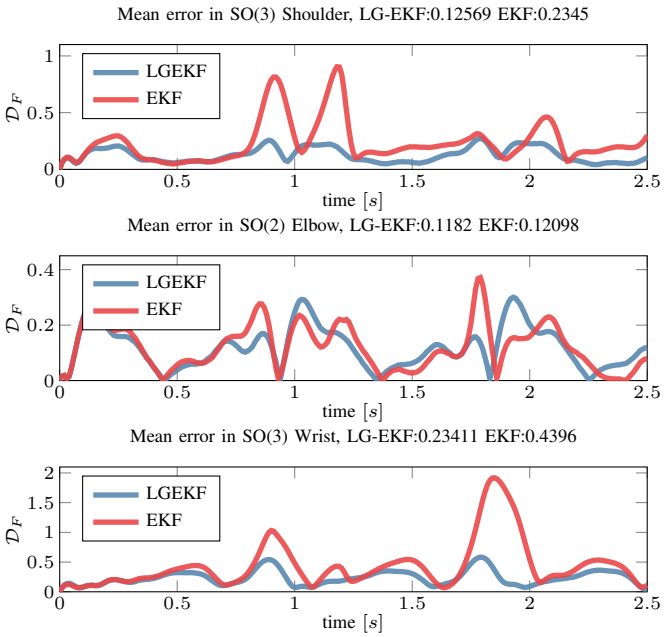

Fig. 5: Rotation matrices error for each of the three joints in the simulated lower body.

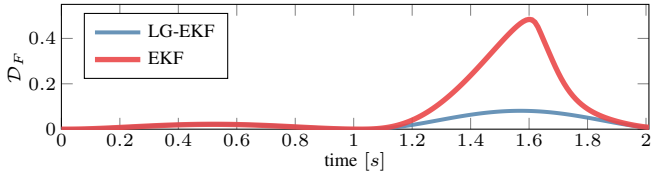

Fig. 6: LG-EKF and EKF estimation during gimbal lock. Both filters accurately estimate the rotation about the y axis until 1 second. After the rotation about y the Euler angle model is in gimbal lock and thus EKF cannot accurately track the orientation until the lock is escaped at 1.5 seconds. LG-EKF estimation is unaffected by gimbal lock.

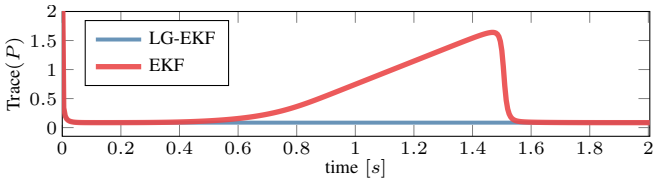

Fig. 7: Trace of the LG-EKF and EKF position error covariance. Both filters start with the same error covariance that quickly converges to a low value. As the Euler angles approach the gimbal lock the EKF position error covariance increases and continues to grow until EKF escapes the lock. LG-EKF position error covariance is unaffected.

effectively puts the Euler angle model into gimbal lock. Next, the model experiences the same 1 s rotation in the now locked world z axis. In order to focus only on the gimbal lock problem, no noise was added to the marker measurements. Measurement noise, process noise, and initial covariances were set as described in Sec. V-A.1.

Figures 6 and 7 show respectively the distance metric described in (30) and the trace of the position error covariance of both filters.

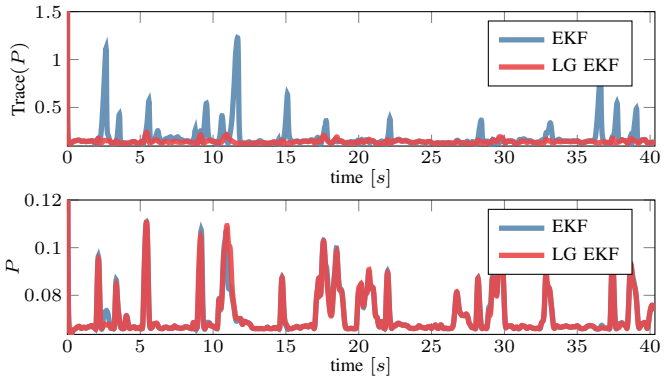

Fig. 8: Position error covariance of LG-EKF and EKF for the spherical shoulder joint (top) and hinge elbow joint (bottom) during boxing motion estimation.

B. Real-world experiment - boxing arm

To evaluate the benefits of estimating real human motion with the proposed method we compare the filters on a highly dynamic boxing motion from the CMU Graphics Lab Motion Capture Database [25]. The movement is captured at 120 Hz with a Vicon motion capture system using 12 cameras. Skeletal model of each participant is created with the Vicon BodyBuilder software and markers are attached at predetermined bony landmarks. We simplify the model by ignoring finger joints and extra joints in the spine Vicon software generates in post processing. In order to focus on the performance of the $\mathsf{SO}(3)$ joint, only the motion of the right arm is estimated. The kinematic chain consists of a spherical joint at the shoulder and a hinge joint at the elbow. Three motion capture markers are used, placed on the upper arm, elbow, and forearm. Figure 3 shows the Euler angle and Lie group models side by side.

To conduct a fair comparison the filters are initialized with the same noise parameters; the initial error covariances, process noise for all states, and observation noise are set to identity, 0.01, and 0.01 respectively. Furthermore, both filters are initialized with a good initial guess obtained from Vicon inverse kinematics available as part of the CMU dataset. We evaluate the performance of each filter by looking at the error covariance as well as using the estimated state to compute the forward kinematics and compare the actual and predicted marker positions. Figure 8 shows the position error covariances of the filters for the spherical shoulder and revolute elbow joints. The shoulder movement is better estimated on the $\mathsf{SO}(3)$ group and thus the error covariance is significantly more uniform than its Euler angles counterpart. Generally, the $\mathsf{SO}(2)$ is expected to behave identically as wrapped $\mathbb{R}^1$ [26]. Table I shows the RMSE between the actual and estimated marker positions. LG-EKF has a better representation of error covariance and avoids gimbal lock at the SO(3) shoulder joint leading to a lower RMSE in the upper arm and elbow markers. The better estimation at the shoulder is propagated through the kinematic chain leading to a lower RMSE in the forearm marker even though the SO(2) joint behaves identically to a single Euler angle.

TABLE I: Root mean squared error in cm between actual and predicted marker positions for boxing arm motion. Where UPA, ELB, and FRA are the upper arm, elbow, and forearm markers respectively. On average LG-EKF improves estimation by 14%.

	UPA	ELB	FRA
EKF	2.61	3.04	2.79
LG-EKF	**2.30**	**2.69**	**2.27**

TABLE II: Root mean squared error in cm between actual and predicted marker positions for markers attached to the waist of the full body model. Where RF, LF, RB, and LB are the right and left, front and back markers respectively. On average estimation on SE3 improves RMSE by 8.2%.

	RF	LF	RB	LB
EKF	1.76	1.91	1.57	1.61
LG-EKF	**1.66**	**1.70**	**1.42**	**1.46**

C. Real-world experiment - full body

To enable localization of the actor in the world frame we add $\mathsf{SE}(3)$ as the first element of LG-EKF's state vector and express the entire full body as a collection of $\mathsf{SO}(3)$ and $\mathsf{SO}(2)$ elements presented in Fig. 1 (right). $\mathsf{SE}(3)$ element connects the world frame to the base of the kinematic model. Shoulders, hips, and neck joints are modelled as $\mathsf{SO}(3)$ elements. Elbows, knees, and wrists are described using a single $\mathsf{SO}(2)$ element and the ankles with two perpendicular $\mathsf{SO}(2)$ elements. A total of 37 markers are attached to the body following the Vicon motion capture manual [25]. To demonstrate the benefits of the $\mathsf{SE}(3)$ representation of localization over a sequence of prismatic and revolute joints and the overall improvement of LG-EKF we use a dynamic full body martial arts movement sequence from the CMU database. Both filters are initialized identically with the same noise parameters as described in Sec. V-B and with a good initial state from the Vicon IK.

Figure 9 compares the position error covariance of the LG-EKF's $\mathsf{SE}(3)$ element state and the EKF's prismatic and revolute joint states. As seen from the uniform covariance, the fast full body rotations and translations are better represented on the $\mathsf{SE}(3)$ group. This can also be observed in the RMSE of the predicted and actual marker positions of the 4 pelvis markers shown in Table II. As an extra comparison we use the Vicon CMU IK results and their more complex full-body model to run forward kinematics and compare the RMSE of predicted and actual marker positions. Table III provides RMSE for the rest of the markers on the body including that of Vicon IK. Even without tuning the noise parameters and initial covariances, the stochastic filtering approaches significantly outperform the Vicon IK method. Furthermore, the LG-EKF achieves a much lower RMSE in almost all the markers over the EKF. The lower error covariance and avoidance of gimbal lock at the SE(3) joint provides a better estimation of the entire skeleton position and orientation. The improvement in the estimation at the base and each SO(3) joint is propagated down the kinematic tree reducing the RMSE of the markers.

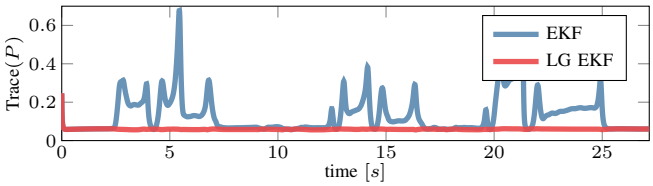

Fig. 9: LG-EKF and EKF position error covariance of the transformation of an $\mathsf{SE}(3)$ from world to base of the kinematic model. Since the transformation is an $\mathsf{SE}(3)$ element LG-EKF is able to accurately estimate it and its evolution over time. Prismatic joints and Euler angles do not correctly represent $\mathsf{SE}(3)$ thus EKF covariance increases during highly dynamic motion.

TABLE III: Root mean square error between predicted and actual marker positions for full body motion capture. LG-EKF outperforms both EKF and VICON IK for most of the markers. Refer to [25] for marker placement and naming details. Note, VICON IK prioritizes ankle markers to avoid unrealistic sliding at the feet.

R Arm	RSHO	RELB	RUPA	RFRM	RWRA	RWRB	
VICON	6.33	4.57	5.17	4.64	7.01	6.91	
EKF	2.89	3.04	**2.82**	2.87	2.43	2.39	
LG-EKF	**2.6**	**2.67**	2.91	**2.58**	**2.29**	**2.23**	
L Arm	LSHO	LELB	LUPA	LFRM	LWRA	LWRB	
VICON	7.82	5.96	6.32	8.19	11.11	10.74	
EKF	2.98	4.51	3.95	2.77	4.22	2.32	
LG-EKF	**2.82**	**4.15**	**3.86**	**2.59**	**4.1**	**2.02**	
Torso and Head	CLAV	T10	STRN	RFHD	LFHD	RBHD	LBHD
VICON	6.09	2.98	2.22	12.85	13.07	10.65	10.59
EKF	1.74	1.55	1.72	1.3	1.18	1.5	1.49
LG-EKF	**1.64**	**1.45**	**1.59**	**1.26**	**1.13**	**1.45**	**1.43**
R Leg	RTHI	RKNE	RSHN	RANK	RHEE	RTOE	RMT5
VICON	3.99	4.78	4.27	**0.4**	1.47	2.54	1.81
EKF	2.06	2.42	2.34	1.15	1.18	0.94	1.06
LG-EKF	**1.93**	**2.4**	**2.33**	1.14	**1.16**	**0.93**	**1.04**
L Leg	LTHI	LKNE	LSHN	LANK	LHEE	LTOE	LMT5
VICON	4.36	4.45	2.48	**0.53**	1.4	2.29	2.4
EKF	2.09	2.01	1.35	1.06	1.22	1	1.18
LG-EKF	**2.07**	**1.98**	**1.34**	1.04	**1.21**	1	**1.16**

VI. Conclusion

We proposed a novel algorithm for human motion estimation based on body worn marker position measurements. The human joints were described as Lie group members, including special orthogonal groups $\mathsf{SO}(2)$ and $\mathsf{SO}(3)$, and a special euclidean group $\mathsf{SE}(3)$. For stochastic inference on Lie groups the LG-EKF was employed, thus explicitly accounting for the non-euclidean geometry of the state space. A constant acceleration motion model for human motion estimation on the group was developed and the Jacobian of the marker position measurements was derived. The performance of the proposed method was evaluated on both simulation and real-world motion capture data, comparing it with the Euler angles-based EKF as well as Vicon IK for full body estimation. We showed that LG-EKF improves estimation for highly dynamic motions and is not affected by gimbal lock.

References

[1] D. Kulić, G. Venture, K. Yamane, E. Demircan, I. Mizuuchi, and K. Mombaur, "Anthropomorphic Movement Analysis and Synthesis : A Survey of Methods and Applications," *IEEE Transactions on Robotics*, vol. 32, no. 4, pp. 776–795, 2016.

[2] T. Sugihara, "Solvability-unconcerned inverse kinematics by the levenberg–marquardt method," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 984–991, 2011.

[3] J. Steinbring, C. Mandery, N. Vahrenkamp, T. Asfour, and U. D. Hanebeck, "High-accuracy real-time whole-body human motion tracking based on constrained nonlinear kalman filtering," *arXiv preprint arXiv:1511.04278*, 2015.

[4] A. Aristidou and J. Lasenby, "Real-time marker prediction and cor estimation in optical motion capture," *The Visual Computer*, vol. 29, no. 1, pp. 7–26, 2013.

[5] J. Steinbring, C. Mandery, F. Pfaff, F. Faion, T. Asfour, and U. D. Hanebeck, "Real-Time Whole-Body Human Motion Tracking Based on Unlabeled Markers," in *IEEE Int. Conf. on Multisensor Fusion and Integration for Intell. Systems (MFI)*. IEEE, 2016, pp. 607–614.

[6] V. Bonnet, G. Daune, V. Joukov, R. Dumas, P. Fraisse, D. Kulić, A. Seilles, S. Andary, and G. Venture, "A constrained extended kalman filter for dynamically consistent inverse kinematics and inertial parameters identification," in *International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2016, pp. 952–957.

[7] T. D. Barfoot and P. T. Furgale, "Associating Uncertainty With Three-Dimensional Poses for Use in Estimation Problems," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679–693, jun 2014.

[8] M. Benallegue and F. Lamiraux, "Estimation and Stabilization of Humanoid Flexibility Deformation Using Only Inertial Measurement Units and Contact Information," *International Journal of Humanoid Robotics*, vol. 12, no. 3, p. 1550025, 2015.

[9] J. Ćesić, I. Marković, I. Cvišić, and I. Petrović, "Radar and stereo vision fusion for multitarget tracking on the special euclidean group," *Robotics and Autonomous Systems*, vol. 83, pp. 338 – 348, 2016.

[10] G. Bourmaud, R. Mégret, M. Arnaudon, and A. Giremus, "Continuous-Discrete Extended Kalman Filter on Matrix Lie Groups Using Concentrated Gaussian Distributions," *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 209–228, 2015.

[11] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating Generic Sensor Fusion Algorithms with Sound State Representations through Encapsulation of Manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57–77, jul 2013.

[12] Q. Rentmeesters, P. A. Absil, P. Van Dooren, K. Gallivan, and A. Srivastava, "An efficient particle filtering technique on the grassmann manifold," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010, pp. 3838–3841.

[13] J. Martinez-Del-Rincon, M. Lewandowski, J. C. Nebel, and D. Makris, "Generalized Laplacian eigenmaps for modeling and tracking human motions," *IEEE Trans. on Cyb.*, vol. 44, no. 9, pp. 1646–1660, 2014.

[14] M. Ding and G. Fan, "Multilayer Joint Gait-Pose Manifolds for Human Gait Motion Modeling," *IEEE Transactions on Cybernetics*, vol. 45, no. 11, pp. 2413–2424, 2015.

[15] M. Devanne, H. Wannous, S. Berretti, P. Pala, M. Daoudi, and A. Del Bimbo, "3-D Human Action Recognition by Shape Analysis of Motion Trajectories on Riemannian Manifold," *IEEE Transactions on Cybernetics*, vol. 45, no. 7, pp. 1340–1352, 2015.

[16] S. Brossette, A. Escande, G. Duchemin, B. Chrétien, and A. Kheddar, "Humanoid posture generation on non-Euclidean manifolds," *IEEE-RAS Int. Conf. on Humanoid Robots*, pp. 352–358, 2015.

[17] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li, *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, 2002.

[18] G. S. Chirikjian, *Stochastic Models, Inform. Theory, and Lie Groups, Vol. 2: Analytic Methods and Modern Applications*. Springer, 2012.

[19] J. M. Selig, "Lie Groups and Lie Algebras in Robotics," in *Comput. Noncommutative Algebra and Applications*, 2005, pp. 101–125.

[20] W. Park, Y. Wang, and G. S. Chirikjian, "The Path-of-Probability Algorithm for Steering and Feedback Control of Flexible Needles," *Int. Journal of Robotics Research*, vol. 29, no. 7, pp. 813–830, 2010.

[21] Y. Wang and G. S. Chirikjian, "Nonparametric Second-Order Theory of Error Propagation on Motion Groups," *The International Journal of Robotics Research*, vol. 27, no. 11, pp. 1258–1273, 2008.

[22] V. Joukov, R. D'Souza, and D. Kulić, "Human pose estimation from imperfect sensor data via the extended kalman filter," in *International Symposium on Experimental Robotics (ISER 2016)*. IEEE, 2016.

[23] D. Q. Huynh, "Metrics for 3d rotations: Comparison and analysis," *J. of Math. Imaging and Vision*, vol. 35, no. 2, pp. 155–164, 2009.

[24] F. S. Grassia, "Practical parameterization of rotations using the exponential map," *J. of graphics tools*, vol. 3, no. 3, pp. 29–48, 1998.

[25] F. De la Torre, J. Hodgins, A. Bargteil, X. Martin, J. Macey, A. Collado, and P. Beltran, "Guide to the carnegie mellon university multimodal activity (cmu-mmac) database," *Rob. Inst.*, p. 135, 2008.

[26] I. Marković, J. Ćesić, and I. Petrović, "On wrapping the Kalman filter and estimating with the SO(2) group," *International Conference on Information Fusion (FUSION)*, pp. 1–6, 2016.