Rhythmic EKF for Pose Estimation during Gait

Vladimir Joukov, Vincent Bonnet, Michelle Karg, Gentiane Venture, Dana Kulič

Abstract—Accurate estimation of lower body pose during gait is useful in a wide variety of applications, including design of bipedal walking strategies, active prosthetics, exoskeletons, and physical rehabilitation. In this paper an algorithm is developed to estimate joint kinematics during rhythmic motion such as walking, using inertial measurement units attached at the waist, knees, and ankles. The proposed approach combines the extended Kalman filter with a canonical dynamical system to estimate joint angles, positions, and velocities for 3 dimensional rhythmic lower body movement. The system incrementally learns the rhythmic motion over time, improving the estimate over a regular extended Kalman filter, and segmenting the motion into repetitions. The algorithm is validated in simulation and on real human walking data. It is shown to improve joint acceleration and velocity estimates over regular extended Kalman Filter by 40% and 37% respectively.

I. INTRODUCTION

Biped walking and human motion imitation are one of the key research areas in humanoid robotics. The first step to human motion imitation is accurate measurement of human motion. Inertial Measurement Units (IMU) are low cost and compact sensors that can be worn on the body and taken into any environment enabling motion measurement outside the lab setting. Estimating pose and learning an accurate movement model using IMUs can lead to the development of better robotic exoskeletons and prosthetics by providing real time control strategies [1]. Pose estimation and modeling are also crucial in physical rehabilitation where IMUs are the preferred method for human motion tracking [2].

There have been many previous works focusing on human pose estimation using IMUs. It is possible to simplify the problem by assuming that the motion is limited to a 2D plane and is slow enough to ignore centripetal acceleration. Then accelerometers can be used as inclinometers to measure the incline of each link [3].

To track fast motions it is possible to add gyroscopes which measure the angular velocity of the limbs. However gyroscopes tend to drift and integrating the measurement accumulates error over time [4]. To combat gyro drift Mariani et al. segmented the gait pattern into phases using angular velocity and computed the initial orientation of each link during the motionless phase between heel strike and toe off using accelerometer data and used forward-backward strap-down integration to compute the limb orientation during the swing phase [5]. This approach requires a separate algorithm to segment the patient’s gait into phases and does not run in real time since both the previous and next foot initial orientation must be known to perform strap down integration. Zhou [6] also employed strap down integration to compute the orientation of each link of the human arm. To deal with drift they used Lagrangian optimization with kinematic constraints. The optimization successfully deals with accelerometer drift due to double integration but does not reduce the gyro drift.

Schwartz et al. [7] used a motion capture studio to record joint angles and IMU sensor data while performing preset activities and applied Gaussian Process Regression to learn the general mapping between IMU sensor data and joint angles for each activity and individual separately. To determine what mapping to use when new sensor data is obtained they used a multi-class support vector machine to learn the classification of activities using sensor data. This method performs with an average joint angle error of 5.6 degrees, but does not generalize to the whole space of human motions or to new subjects.

Many of the methods described above treat each link as a separate rigid body and compute their orientations independently without considering human kinematic constraints. The orientation of each rigid body is converted to joint angles in a post processing step that requires optimization and may lead to solutions which are not physically attainable. Seel et al. [8] used optimization to determine sensor transforms with respect to the joints based on an arbitrary calibration motion then computed the joint angle by integrating gyroscope difference and projecting accelerometer readings into the joint plane. Their method achieves an accuracy of 3° for flexion/extension but cannot handle arbitrary 3 dimensional motion. Lin and Kulič [9] used a kinematic model and a Extended Kalman Filter (EKF) to estimate arbitrary 3D leg motion. The states of the EKF are the joint positions, velocities, and accelerations, while the measurement vector consists of the accelerometer and gyroscope data from sensors attached above the knee and ankle. The EKF inherently deals with sensor noise and produces joint estimates directly, which can be used for real time measurement and feedback. In Joukov et al. [10] this approach was extended to model movement when the base is not stationary, such as gait. However, the constant acceleration assumption of the EKF model reduces accuracy at turning points and gyroscope drift integrates into errors in the joint angles estimation.

In this paper we propose an approach to combine the EKF with an online adaptive canonical dynamical system (CDS) [11] to estimate the lower body joint angles for gait and other rhythmic movements. The motion is recorded with five small wearable inertial measurement units (IMUs) and the lower body is described using a kinematic model with

*V. Joukov, M. Karg, and D. Kulič are with the University of Waterloo, Canada. V. Bonnet and G. Venture are with the Tokyo University of Agriculture and Technology, Japan.
three DoFs in the hip and one DoF in the knee, enabling the estimation of arbitrary leg motion in 3D space. The states of the EKF are the positions, velocities, and accelerations of the joints and the measurement vector are the measured 3 dimensional linear accelerations and angular velocities [9]. Taking advantage of the fact that gait and many other human movements are periodic, a CDS is learned online from the joint velocities estimated by the EKF. Combining the EKF with the CDS integrates a model of the periodicity of the movement to improve joint angle estimation. The CDS directly provides a set of required measures such as the frequency and phase of the movement which can be used for segmentation and subsequent motion analysis.

II. PROPOSED APPROACH

An estimate of the lower body pose is calculated in real time using data from several IMUs. The data is combined using an EKF to determine the joint angles of the lower body. A CDS is adapted to the underlying motion, improving the performance of the EKF over time by removing the constant acceleration assumption. The CDS also extracts useful features from the estimated motion, which can be used to evaluate gait properties and isolate gait segments. Figure 1 provides a schematic of the algorithm.

A. Lower Body Kinematics

The human lower body is modeled as a set of rigid bodies connected through revolute joints describing the motions achievable at the pelvis (ball joint), hips (ball joints), and knees (hinge joints). Five IMUs are located at the lower shank, lower thigh and lower trunk levels (Figure 2). To model the translation and rotation of the pelvis, three prismatic and revolute joints connect it to the inertial frame [10].

With the frames assigned we can use forward kinematics to compute the position, velocity, and acceleration of each frame using those of the previous frame recursively. Considering the angular velocity of the previous frame \( w_{i-1}^i \) and the joint velocity \( \dot{q}_i \) caused by actuation of the revolute joint \( i \), the angular velocity of the next frame can be computed as

\[
\dot{w}_i^i = R_{i-1,i}^T w_{i-1}^{i-1} + R_{i-1,i}^T \dot{q}_i \tag{1}
\]

where \( R_{i-1,i} \) is the rotation matrix from frame \( i-1 \) to frame \( i \). Taking the derivative of (1) the angular acceleration is:

\[
\ddot{w}_i^i = R_{i-1,i}^T \dot{w}_{i-1}^{i-1} + R_{i-1,i}^T \ddot{q}_i + \dot{w}_i^i \times (R_{i-1,i}^T \dot{q}_i) \tag{2}
\]

The linear velocity can be calculated using the cross product of angular velocity and the displacement vector \( r_i \) from the previous frame to the next for a revolute joint

\[
x_i = R_{i-1}^T x_{i-1} + R_{i-1,i}^T \dot{q}_i \tag{3}
\]

and directly using joint velocity for prismatic joints.

\[
x_i = R_{i-1}^T x_{i-1} + w_i^i \times r_i \tag{4}
\]

Taking the derivative and adding gravity in the current frame, the linear acceleration in the current frame is:

\[
xdot_i = R_{i-1}^T \ddot{x}_{i-1} + \alpha_i^i \times r_i + w_i^i \times w_i^i \times r_i + R_{0,i}g \tag{5}
\]

for a revolute joint \( i \) and

\[
xdot_i = R_{i-1}^T \ddot{x}_{i-1} + \dot{q}_i + R_{0,i}g \tag{6}
\]

for a prismatic joint \( i \). \( R_{0,i} \) is the rotation from world frame to current frame, \( g \) is the gravity vector. In the knee and thigh frames, the angular velocity vectors \( \omega \) and the linear acceleration vectors \( \ddot{x} \) are measured by the IMUs.

B. Rhythmic Extended Kalman Filter for Kinematic Chain

The Kalman Filter [12] is a popular sensor fusion technique that estimates the state of a system from noisy observations. For a linear model, it is shown to be an optimal filter under the assumption that both measurement and process noise are zero-mean Gaussian. For a non linear system, the Extended Kalman Filter linearizes the equations about the operating point. The equations are approximated as

\[
z_t \approx \tilde{z}_t + C(s_t - \tilde{s}_t) + v_t \tag{7}
\]

\[
s_t \approx \tilde{s}_t + A(s_t - \tilde{s}_t) + w_{t-1} \tag{8}
\]

where \( A \) and \( C \) are the Jacobians of the state update and measurement equations with respect to the state \( s \), \( \tilde{s} \) is the noiseless state estimate, \( \tilde{z} \) is the noiseless measurement estimate, \( v_t \) is the measurement noise, and \( w_{t-1} \) is the process noise. The filter calculates the optimal Kalman gain \( K \) that
minimizes the error co-variance matrix \( P = cov(s_t - \hat{s}_t) \) and at each iteration updates the state as
\[
\hat{s}_t = f(\hat{s}_{t-1}) + K(z_t - h(f(\hat{s}_{t-1})))
\] (9)
Where \( \hat{s}_t \) is the Kalman state estimate at time-step \( t \) [12].

For our purposes, the state vector consists of the position \( q \), velocity \( \dot{q} \), and acceleration \( \ddot{q} \) of the joint angles and the measurement vector includes the 3D acceleration and angular velocity IMU sensor readings. The state at the next time-step is predicted by integrating the velocity and acceleration terms [9], assuming any change in the acceleration is part of the noise:
\[
\begin{align*}
q_t &= q_{t-1} + \dot{q}\Delta t + \ddot{q}\Delta t^2/2 \\
\dot{q}_t &= \dot{q}_{t-1} + \ddot{q}\Delta t \\
\ddot{q}_t &= \dddot{q}_{t-1}
\end{align*}
\] (10-12)
where \( \Delta t \) is the time difference between measurements. However, this leads to lag in the acceleration estimate which propagates to the velocity and position estimates. Instead of a constant acceleration we can add higher order terms such as jerk into the model and assume constant jerk.
\[
\begin{align*}
\dddot{q}_t &= \dddot{q}_{t-1} + \dddot{q}\Delta t \\
\ddot{q}_t &= \ddot{q}_{t-1} + \dddot{q}\Delta t \\
\dot{q}_t &= \dot{q}_{t-1}
\end{align*}
\] (13-14)
Unfortunately IMU sensors cannot measure jerk and thus no extra information is available to improve the acceleration estimate.

Consider the case when the motion is rhythmic, such as human gait. Then we can assume jerk is a function of the phase \( \Omega \) and frequency \( f \) of the motion:
\[
\dddot{q} = F(\Omega)
\] (15)
\[
\dot{\Omega} = f
\] (16)
In this paper, we propose to learn \( F(\Omega) \) during online estimation by learning an analytically differentiable joint velocity model and gradually improve the performance of the EKF over time. We use the CDS to learn the underlying harmonic Fourier series of the joint velocity and use the second analytical derivative as the jerk \( F(\Omega) \).

1) Canonical Dynamical System: The CDS is an estimator based on the harmonic Fourier series
\[
\hat{y}(t) = \sum_{i=0}^{n} a_i \cos(i\Omega) + \sum_{i=1}^{n} b_i \sin(i\Omega)
\] (17)
where \( \hat{y} \) is the state estimate, \( a_i \) and \( b_i \) are the Fourier coefficients, and \( \Omega \) is the current phase. A feedback adaptive frequency phase oscillator is used to adapt the coefficients and phase to learn a rhythmic motion online.
\[
\dot{\Omega} = f - \zeta \cos(\Omega)
\] (18)
\[
\dot{f} = -\zeta \cos(\Omega)
\] (19)
\[
\dot{a}_i = \eta \cos(i\Omega) e
\] (20)
\[
\dot{b}_i = \eta \sin(i\Omega) e
\] (21)
Where \( e \) is the error between actual and estimated values \( e = y - \hat{y} \) and \( \zeta \) and \( \eta \) are the frequency and coefficient learning rates respectively [11].

To learn the rhythmic jerk \( F(\Omega) \) using this method actual jerk values are required which are not available from the EKF estimated state. Noting that the Fourier series are infinitely differentiable the CDS can be used to learn the lower order terms such as acceleration or velocity, and then analytically differentiated to compute \( F(\Omega) \). Starting the Fourier coefficients at zero ensures there will be no initialization problems since EKF will converge to an accurate pose estimate using gravity measurements and guarantees initial performance of the Rhythmic-EKF is identical to that of regular EKF. For pose estimation using IMU sensors we choose joint velocity as the term to learn since it should be less affected by bias, and is not assumed to be constant, as is acceleration before the CDS is converged.

2) Virtual Yaw Sensor: The accelerometer’s constant gravity reference ensures that EKF will accurately correct gyroscope drift coinciding with the world \( x \) and \( y \) axis. However EKF cannot correct gyroscope drift around the world \( z \) axis using the accelerometer data, this can result in accumulation of error in joint angle estimates for joints which rotate about axes parallel to the direction of gravity. To handle this issue, Lin and Kulic proposed applying potential fields in the EKF joint acceleration state once the joint position exceeds the joint limit [9]. The approach requires the potential field to be tuned for each joint independently and may cause oscillation of position estimate between the joint limits. Instead we use a virtual yaw sensor \( \gamma \) as part of the measurement vector. It requires only a single tuning parameter and effectively combats gyro drift.

We assume that, for the most part, the human is walking in a straight line and thus the yaw angles of their pelvis and each thigh should not drift away from the original orientation. With this in mind we attach a virtual yaw sensor to the pelvis and each thigh, the measurement is set to the starting orientation and is added as part of EKF measurement vector. The measurement noise of the virtual yaw sensor represents the standard deviation of expected motion and is used as a tuning parameter. To allow for curved walking and turning the pelvis yaw sensor can be removed and the thigh sensors can be set to measure yaw with respect to the pelvis frame instead of the world frame.

Using forward kinematics the yaw measurement prediction is computed from entries of the rotation matrix as:
\[
\gamma = \tan^{-1}\left(\frac{R_{0,i}(2,1)}{R_{0,i}(1,1)}\right)
\] (22)
and the Jacobian of the yaw sensor is the third row of the velocity Jacobian [13] in the pelvis frame. Figure 3 shows the benefits of using a virtual yaw sensor in EKF when the gyroscope measurement is biased. Note that the 3D motion is still observable and estimated, but the yaw sensor prevents the angle from drifting with time.
III. RESULTS

The approach is evaluated in simulation and with human gait data. In simulation, the convergence of the CDS and its impact on EKF estimation are analyzed with Gaussian sensor noise. Next the Rhythmic-EKF is compared to optical motion capture on human marching data.

A. Simulation Results

To test the convergence properties of Rhythmic EKF we simulate a gyroscope and an accelerometer attached at an offset of 0.5m to a single revolute joint actuating about world z axis. To simulate human motion the joint is actuated for 100 seconds using a Fourier series with 10 harmonics starting at 4 rads/s and coefficients from a uni-variate distribution. Using forward kinematics, the sensor measurements are computed and random Gaussian noise of $2^5 m^2$ and $0.5 rad/s$ is added to the accelerometer and gyroscope respectively. Next the EKF’s noise parameters are found using Matlab’s constrained minimization toolbox, by finding the noise parameters that minimize the joint position estimation error. The Rhythmic-EKF uses the same noise parameters as the EKF, the number of harmonics and well as frequency and coefficient learning rates were chosen experimentally to be 7, 0.7, and 0.2 respectively. We verify that the Rhythmic-EKF can successfully adapt the CDS to the rhythmic motion and compare the performance to EKF after convergence.

Figure 4 shows the adaptation process of Rhythmic-EKF to the rhythmic motion. Frequency is accurately tracked after 10 repetitions and the Fourier coefficients are learned within 20 repetitions. Once convergence is achieved the tracking is improved for all states, i.e., the acceleration, position, and velocity over EKF with optimal noise parameters. Particularly lag, overshoot and undershoot in acceleration and velocity estimates are reduced. Furthermore the predicted phase can easily be used for segmentation into repetitions and frequency can be used to track the speed of the motion. A closer view of one of the velocity peaks is shown in figure 5. Since acceleration is no longer assumed constant the predicted velocity does not overshoot or significantly lag the actual value.

EKF’s acceleration estimate error increases linearly with jerk since the model always predicts constant acceleration, Rhythmic-EKF builds an accurate model over time. Thus by plotting jerk against error in acceleration estimate (figure 6) the benefits of Rhythmic-EKF are very clear.
link lengths to generate the kinematic model. The hip joint centers were determined using the pelvis width and length as well as leg length as described in Harrington et al. [15]. The position of the knee and ankle joint centers were obtained by taking the average of their respective medial and lateral marker positions. Data from the sensors and the motion capture was time-aligned in post processing.

Joint angles, velocities, and accelerations were obtained from marker data. The proposed Rhythmic-EKF and regular EKF algorithms were applied to the IMU data. Three virtual yaw sensors were added to the measurement vector at the center of the pelvis and each thigh to combat gyroscope drift. The noise parameters were tuned for best performance of the EKF algorithm for the first participant. Rhythmic-EKF was set to have 7 harmonics and the frequency and coefficient learning rates were set to 0.7 and 0.05 respectively. To phase synchronize all of the joints we use \( f \) and \( \Omega \) estimated for the right knee joint for the entire lower body. First the metronome trials were used to verify convergence to accurate frequency and phase. Figure 7 shows Rhythmic-EKF convergence to human motion over one of the trials. Table I shows the accuracy of the frequency convergence after the first 15 seconds of walking. At initialization the Fourier coefficients of the CDS are zero resulting in zero estimated jerk and Rhythmic-EKF behaving identically to regular EKF. Once the frequency is locked on the Fourier coefficients begin to converge and start improving the estimate of jerk.

Next the non-metronome data was used to compare Rhythmic-EKF and EKF. The two algorithms are expected to have the same performance during constant acceleration regions. Using data-sets of participants walking at their own pace, we verify that the proposed approach improves estimation over regular EKF in high acceleration regions.

The Rhythmic-EKF model also significantly improves measurement prediction which can be useful in control applications. Table III shows the RMSE between actual and predicted gyroscope and accelerometer measurements for the next time step for the knee and ankle IMU.

Table II presents the hip and knee joint position, velocity, and acceleration root mean squared error (RMSE) separately into regions where acceleration is below or above 75% of maximum. Even without noise parameters tuned for Rhythmic-EKF, on average, it improves the velocity and acceleration estimation by 35% and 55% respectively in high acceleration regions. In low acceleration regions Rhythmic-EKF still outperforms EKF but the difference is not as drastic in acceleration estimation. The results in bold indicate the best performing system for the corresponding row. On average Rhythmic-EKF improves acceleration and velocity estimation over EKF by 40% and 37% respectively. It outperforms EKF and other IMU based pose estimation methods [8], [7] with an average RMSE of 2.4°.

The Rhythmic-EKF model significantly improves measurement prediction which can be useful in control applications. Table III shows the RMSE between actual and predicted gyroscope and accelerometer measurements for the next time step for the knee and ankle IMU.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Actual</th>
<th>Rhythmic-EKF</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subj1</td>
<td>1.167</td>
<td>1.170</td>
<td>-0.30</td>
</tr>
<tr>
<td>Subj2</td>
<td>1.167</td>
<td>1.179</td>
<td>-1.08</td>
</tr>
<tr>
<td>Subj3</td>
<td>1.167</td>
<td>1.169</td>
<td>-0.18</td>
</tr>
</tbody>
</table>

---

**TABLE I** Mean frequency after the initial 15 seconds of walking estimated by the Rhythmic-EKF.

<table>
<thead>
<tr>
<th></th>
<th>Rhythm. EKF</th>
<th>EKF</th>
<th>% Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bottom 75% Acceleration Region</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Joint Position RMSE (deg/s)</td>
<td>103.9</td>
<td>105.4</td>
<td>193.8</td>
</tr>
<tr>
<td>Knees Joint Position RMSE (deg/s)</td>
<td>162.5</td>
<td>162.2</td>
<td>339.3</td>
</tr>
<tr>
<td><strong>Top 25% Acceleration Region</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hip Joint Position RMSE (deg/s)</td>
<td>6.79</td>
<td>6.48</td>
<td>35.00</td>
</tr>
<tr>
<td>Knees Joint Position RMSE (deg/s)</td>
<td>24.77</td>
<td>24.82</td>
<td>58.72</td>
</tr>
</tbody>
</table>

**TABLE II**: Root mean squared error of joint position, velocity, and acceleration for hip and knee joints (averaged over left and right) over bottom 75% (left) and top 25% (right) acceleration regions with participants marching at their own pace.

<table>
<thead>
<tr>
<th></th>
<th>Rhythm. EKF</th>
<th>EKF</th>
<th>Rhythm. EKF</th>
<th>EKF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knee IMU</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gyroscope RMSE (deg/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subj1</td>
<td>10.08</td>
<td>11.55</td>
<td>17.52</td>
<td>21.23</td>
</tr>
<tr>
<td>Subj2</td>
<td>10.69</td>
<td>13.05</td>
<td>15.51</td>
<td>20.45</td>
</tr>
<tr>
<td>Subj3</td>
<td>10.98</td>
<td>10.82</td>
<td>18.78</td>
<td>23.08</td>
</tr>
<tr>
<td><strong>Ankle IMU</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accelerometer RMSE (m/s²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subj1</td>
<td>1.18</td>
<td>0.96</td>
<td>1.35</td>
<td>1.18</td>
</tr>
<tr>
<td>Subj2</td>
<td>1.18</td>
<td>1.05</td>
<td>1.30</td>
<td>1.24</td>
</tr>
<tr>
<td>Subj3</td>
<td>1.21</td>
<td>1.15</td>
<td>1.34</td>
<td>1.48</td>
</tr>
</tbody>
</table>

---

Once Rhythmic-EKF achieves frequency convergence the
estimated phase can be used for segmenting the motion and temporally aligning the segments. We allow the algorithm to run for 15 seconds before using the phase variable for segmentation. Between the 3 participants a total of 306 steps were taken, there were 11 visibly incorrect segments in the hip joints and 7 in the knee joints leading to an accuracy of 96.4% and 97.1% respectively. Figure 9 shows the right knee joint angle estimate segmented using the phase variable.

Due to quick convergence of the frequency and phase variables Rhythmic-EKF can handle changing frequencies. A metronome walking trial was collected and the metronome frequency was changed every 30 seconds. The algorithm was able to accurately track the frequency and phase of the motion. The phase was used for segmentation and temporal alignment. Figure 10 shows tracking of the frequency and phase and motion segmentation based on the estimated phase.

With the motion segmented and the segments aligned, important measures such as mean and standard deviation of the joint angles can be extracted. Because all of the lower body joints are synchronized to a single phase variable, it is easy to visualize the symmetry between the left and right sides which is very useful in physiotherapy. The learned symmetry can also be used to develop control strategies for single side rehabilitation exoskeletons. Figure 11 shows the mean and standard deviation of hip and knee joint angles during the marching exercise.

IV. CONCLUSION

This paper proposes a rhythmic version of the extended Kalman filter for lower body pose estimation using inertial measurement units. The algorithm learns the underlying rhythmic motion incrementally during observation, improves the pose estimate over regular Kalman filter, and segments the motion into repetitions. With added virtual yaw sensors to combat gyro drift the system achieves an average RMSE of 2.4 degrees in the knee and hip joints during walking.

Future work will include using the estimated phase of the rhythmic motion to detect heel strikes and switch the base of the kinematic model.

REFERENCES