

Predictive trajectory estimation during rehabilitative tasks in augmented reality using inertial sensors

Christopher L. Hunt*, Avinash Sharma*, Luke E. Osborn*, Rahul R. Kaliki†, and Nitish V. Thakor*‡

*Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218 USA

†Infinite Biomedical Technologies, LLC, Baltimore, MD 21218 USA

‡Singapore Institute for Neurotechnology, National University of Singapore, 119077 Singapore

Email: chunt11@jhmi.edu

Abstract—This paper presents a wireless kinematic tracking framework used for biomechanical analysis during rehabilitative tasks in augmented and virtual reality. The framework uses low-cost inertial measurement units and exploits the rigid connections of the human skeletal system to provide egocentric position estimates of joints to centimeter accuracy. On-board sensor fusion combines information from three-axis accelerometers, gyroscopes, and magnetometers to provide robust estimates in real-time. Sensor precision and accuracy were validated using the root mean square error of estimated joint angles against ground truth goniometer measurements. The sensor network produced a mean estimate accuracy of 2.81° with 1.06° precision, resulting in a maximum hand tracking error of 7.06 cm. As an application, the network is used to collect kinematic information from an unconstrained object manipulation task in augmented reality, from which dynamic movement primitives are extracted to characterize natural task completion in $N = 3$ able-bodied human subjects. These primitives are then leveraged for trajectory estimation in both a generalized and a subject-specific scheme resulting in 0.187 cm and 0.161 cm regression accuracy, respectively. Our proposed kinematic tracking network is wireless, accurate, and especially useful for predicting voluntary actuation in virtual and augmented reality applications.

I. INTRODUCTION

In recent years, augmented and virtual reality environments have gained popularity as viable techniques for use in rehabilitation technologies due to their flexibility and ease of deployment. There have been studies exploring the efficacy of such techniques in the space of post-stroke rehabilitation [1], myoelectric prosthesis control [2], and brain-computer interfaces [3]. In such studies, it is often the task to have subjects achieve some kinematic goal (e.g. picking and placing an object) in the virtual space while the dynamics and mechanics of the solution are measured and analyzed. However, a major issue still prevalent in these rehabilitation technologies is the ability to actuate naturally in the virtual world. Common sensors used to capture physical motion and translate it into virtual motion include optical kinematic trackers [4], depth motion cameras [5], and electrogoniometers [6]. While optics-based trackers are considered state-of-the-art due to their high tracking fidelity, they are both expensive and sensitive to movements that result in partial visual occlusion of the target. Additionally, while goniometers are common in classical physical rehabilitation systems, they are limited to single joint measurements and are only practical for static analysis.

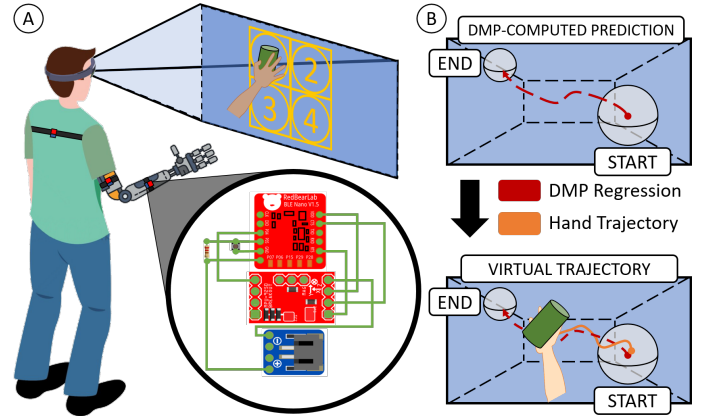


Fig. 1. An overview of a rehabilitation session. (A) The individual uses an augmented reality headset to receive kinematic tasks to complete. Tasks consist of transporting an object to and from different quadrants while possibly changing its orientation. Sensorized tracking nodes are rigidly affixed to the anatomical landmarks defined in [7] and are used to record multi-joint trajectories for primitive construction. (B) Once computed, these primitives are used to predict natural, user-specific hand trajectories in subsequent tasks. These predicted trajectories can then be rendered by the headset to serve as an optimal reference for the user.

In response to some of these limitations, inertial measurement units (IMUs) have become popular actuators in human-machine interfaces that require arm tracking capabilities [8]. An IMU is a microelectromechanical system that reports device orientation through the fusion of linear acceleration, angular velocity, and magnetic field sensor measurements. Multiple computational techniques have been designed to fuse these raw sensor values into representations of device orientation including variations of Kalman filtering and constrained estimation [9], [10].

In this work, we present a wireless kinematic tracking framework that is suitable for rehabilitation tasks in augmented and virtual reality due to its low cost, low weight, and high accuracy. Furthermore, we demonstrate how kinematic information gathered from tasks in augmented space can be used to extract movement primitives useful for individual trajectory modeling.

II. METHODS

The sensor network consists of a single hub with up to eight sensorized nodes in a star topology. Each node in the network

consists of a nRF51822 microcontroller (RedBearLab, Shenzhen, China) with a MPU9250 9-axis IMU (InvenSense, San Jose, CA). Measurements from the accelerometer, gyroscope, and magnetometer are fused using an on-board implementation of the Mahony complementary filter [11] to provide a quaternion-based representation of the device's orientation relative to the Earth's magnetic field. These nodal orientation estimates are then transmitted back to the central hub for further processing using the Nordic Enhanced ShockBurst protocol. This proprietary protocol was chosen over traditional standards, such as Bluetooth Low Energy or Wi-Fi, due to its superior transaction latency (0.5 ms) and low power consumption [12].

Before the network can be used for functional arm tracking, subjects must complete a short session calibration routine each time the nodal sensors are donned. The goal of the routine is to measure and negate the orientation offset from the skeletal segment's center of rotation. To accomplish this, we follow the calibration routine outline in [13]. Using this calibration method, no significant drift in orientation estimates was observed during use (typical session time: 1 hour).

Once calibrated, orientation estimates are aggregated at the central hub and skeletal joint angles and positions can be estimated through quaternion manipulation. For a specific joint j that lies between devices with orientations q_i and q_{i+1} , we can compute the angle of j by computing the relative rotation that maps q_i to q_{i+1} . Assuming q_i and q_{i+1} are both normalized, this rotation can be computed with equation (1), where inversion implies the normalized conjugate quaternion and multiplication implies the Hamilton product [14].

$$q_j = q_i^{-1} q_{i+1} \quad (1)$$

Given q_j , the joint angle θ_j can then be calculated with equation (2), where $q_{j,0}$ denotes the scalar part of q_j .

$$\theta_j = 2 \cos^{-1} q_{j,0} \quad (2)$$

With this information, three-dimensional joint positions can be estimated using a compounding iterative estimation algorithm (see Algorithm 1). In this procedure, we first compute the relative rotation q_j for a specific joint. We then rotate a vector v_i that represents the translation that maps from joint j to joint $j + 1$ when in the calibration configuration. This rotation can be expressed as the following operation:

$$v'_i = q_j v_i q_j^{-1} \quad (3)$$

This rotated vector, v'_i , represents the Euclidean position of joint j relative to joint $j - 1$. To obtain this position relative to the global egocentric origin, we simply add this estimate to the position estimate of the previous joint $j - 1$.

TABLE I
OBJECT MANIPULATION TASKS

Task No.	1	2	3	4	5	6	7	8
Start Config.	3H	3V	4V	4H	1V	1V	1V	2V
Final Config.	4H	1V	3H	2V	2V	2H	4V	1H

Algorithm 1: Estimate Euclidean position of joints

Input: A list of N device orientations, q
A list of $N - 1$ limb vectors, v
Output: A list of N joint position estimates, p
 $p \leftarrow \{\vec{0}\}$
for $i \leftarrow 0$ **to** $N - 1$ **by** 1 **do**
 $q_j \leftarrow q_i^{-1} q_{i+1}$
 $p'_i \leftarrow q_j v_i q_j^{-1}$
 $p \leftarrow p \cup (p_i + p'_i)$
end

The kinematic tracking framework thus far described has been used to characterize motor control in rehabilitation tasks performed in augmented reality using the Microsoft HoloLens (Microsoft Corp., Redmond, WA). $N = 3$ able-bodied subjects were asked to complete a series of pick-and-place tasks in augmented reality modeled after the Prosthesis Hand Assessment Measure [7]. Subjects were instructed to manipulate a cylindrical object in augmented space in as natural a manner as possible (see Fig. 1). Tasks are defined by their initial and final positions (1-4) and orientations (horizontal or vertical). Tasks definitions can be found in Table I, however, readers are referred to [15] for further implementation details. Three sensor nodes used for tracking were rigidly affixed to the torso, upper arm, and forearm using Velcro® fasteners. Once collected, the task kinematics were then used in a leave-one-out (LOO) scheme to generate movement models that were then used to predict trial trajectories from unseen data. All experiments were approved by the Johns Hopkins Medicine Institutional Review Board.

III. RESULTS AND DISCUSSION

A. Sensor Validation

Sensor accuracy and precision are validated using a digital goniometer fashioned from a high-precision stepper motor with a 0.9° step size (NEMA, Rosslyn, VA). A sensor node was rigidly affixed to the motor shaft while the motor was driven periodically with 1/8 microstepping. Shaft angle was then estimated independently using equation (2) and compared to the ground truth from the internal encoder.

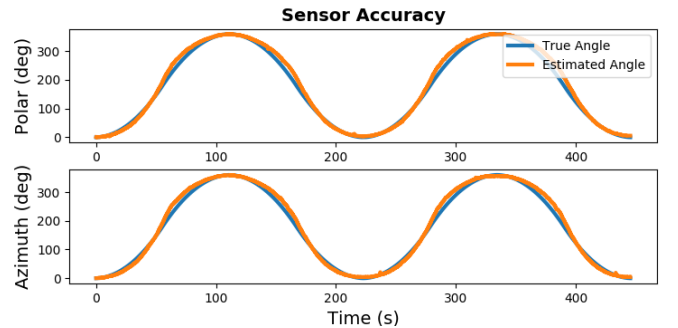


Fig. 2. The angle estimation error over two periods in two sensor configurations: polar and azimuthal. Estimated and true angle values are shown in orange and blue respectively. Sensor accuracy is 3.38° (1.34° precision) and 2.24° (0.77° precision) for the polar and azimuthal configurations.

ABLE-1 Hand Trajectory DMPs For Pick-and-Place Tasks

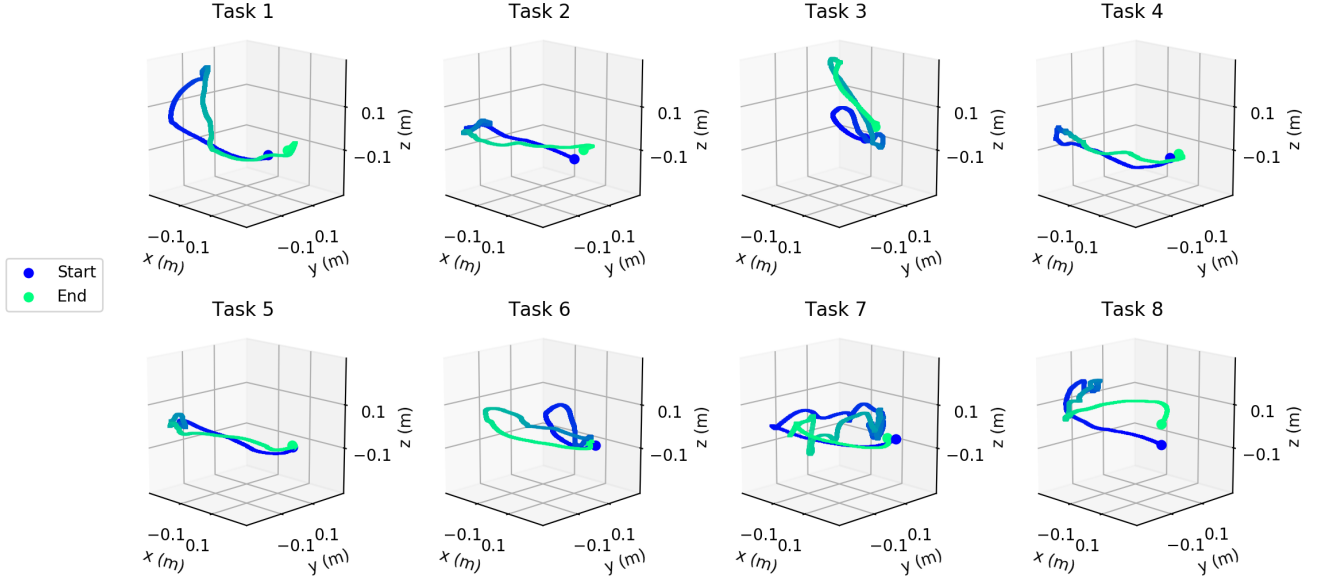


Fig. 3. An example of task-specific DMPs computed from the time-normalized kinematics of a single subject (ABLE-1). These DMPs were calculated as the mean trajectory (in meters) followed across trials. DMPs were smoothed by applying a Savitzky-Golay filter with a window size equal to 5% of the total signal length. For visualization, hand trajectories progress temporally from start (blue) to end (green).

Accuracy is defined as the root mean square error (RMSE) of the estimated and true shaft angle while precision is defined as the median variance of the estimated shaft angle at corresponding timestamps across periods. This validation was performed with the nodal sensor in both polar (horizontal) and azimuthal (vertical) configurations. For the polar configuration, the sensor showed an accuracy of 3.38° with 1.34° precision while these values are 2.24° and 0.77° for the azimuthal configuration. These results can be seen in Fig. 2.

Given a maximum azimuthal angular deviation of $d\theta$ and polar angular deviation of $d\phi$, the maximum translational deviation is the L_2 -norm of the differential displacement in spherical coordinates [16]:

$$\epsilon = \sqrt{dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2)} \quad (4)$$

If we assume a constant radius of rotation ($dr = 0$) and maximal contribution from the polar deviation ($\sin \theta = 1$), the expression for ϵ reduces to:

$$\epsilon = \sqrt{r^2 (d\theta^2 + d\phi^2)} \quad (5)$$

For $d\theta = 3.37^\circ$ and $d\phi = 2.24^\circ$, $\epsilon = 7.06$ cm for a 1 m radius of rotation. For a multi-joint rigid body, the overall translational deviation at the end-effector is the linear summation of ϵ_j at each joint j relative to joint $j - 1$. Therefore, ϵ serves as a maximum on the hand-tracking displacement error as the average human has an arm length much less than 1 m. While this tracking error is greater than current start-of-the-art optical tracking methodologies, it is comparable to depth-based optical sensors currently widespread in the fields of rehabilitation and robotics. For example, the Microsoft Kinect 2 achieves a mean tracking accuracy of 9.3 ± 1.56 cm in

standing rehabilitation tasks [17]. In addition to slightly better performance, this system is much more flexible than traditional optical systems because it is not susceptible to movement occlusion or environmental illumination and, therefore, can be used during truly unconstrained motion.

B. Predictive Trajectory Estimation

Unnormalized kinematics were collected for three trials of eight tasks each and from these kinematics, end-effector trajectories were estimated using Algorithm 1. These estimated trajectories are then used to define dynamic movement primitives (DMPs) that characterize the motion for a task. These task-specific DMPs were computed by averaging the time-normalized task trajectories across each trial and applying a smoothing Savitzky-Golay filter with a 5% window length. In this way, each task defines a single DMP to be used for later trajectory prediction.

These task-specific DMPs (see Fig. 3) were used to predict end-effector trajectories in a LOO analysis. This was accomplished by defining the transition function of an unscented Kalman filter as the dynamics of the DMP [18]. To validate the efficacy of such prediction methods, the RMSE between the predicted and actual task trajectories are computed in two schemes. In the first method, DMPs were averaged over all subjects and used to estimate trajectories for all subjects. This yielded an average regression accuracy of 0.187 cm. However, this method is sensitive to inter-subject variability which can result in very poor regression results. Fig. 4 is an example where the generalized DMP regression performs significantly worse than the average.

In the second method, task DMPs are kept subject-specific and are used to predict only the remaining trial trajectory

DMP-Based Trajectory Regression

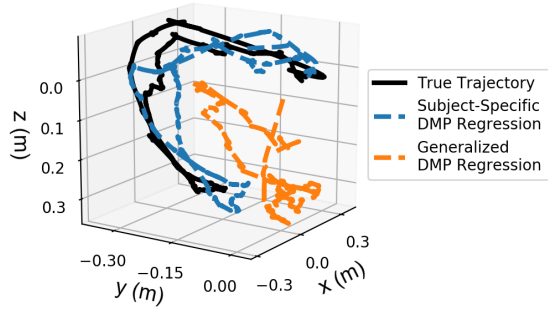


Fig. 4. Estimated trajectory from a single subject computed using the generalized (orange) and subject-specific (blue) DMP model to inform the transition function of the unscented Kalman filter.

not used for DMP construction. Compared to the generalized DMPs, these subject specific models perform better at trajectory prediction with an average accuracy of 0.161 cm. This trend is present in all manipulation tasks, although this is not surprising as each individual constructs a unique motor plan for task completion when unconstrained [19]. On the other hand, regression for Tasks 7 and 8 showed very little performance difference when comparing the generalized and subject-specific methods, 0.60% and 0.22% percent change respectively. This suggests a universality in the motion plan individuals compute for these specific manipulations, however, more subjects are needed to explore this further.

IV. CONCLUSION

In this work, we have presented a kinematic tracking framework that leverages IMUs and rigid body constraints to accurately estimate joint position to the centimeter range. Furthermore, we demonstrate the efficacy of the network through both a maximum error analysis and a functional application: a pick-and-place rehabilitation task in augmented reality. We show that kinematic data gathered from tasks such as these can be used to extract DMPs for predictive trajectory estimation. This system has applications for virtual training and motor analysis of patients suffering from motor disabilities such as stroke, paresis, and limb amputation and future extensions of this work would like to apply these methods of regression on upper-limb amputees in particular. Because the results of the regression analysis suggest that internal models for kinematic task completion are consistent within a subject, we would like to ascertain these primitives' predictive power when trained on a healthy limb but assessed on the artificial limb of an amputee.

ACKNOWLEDGMENT

The authors would like to thank the human subjects who participated in this study; Infinite Biomedical Technologies; the Applied Physics Laboratory; the National Institutes of Health; and The Johns Hopkins University. This was supported in part by the National Institutes of Health under Grants No. T32EB00338312 and R44HD072668.

Trajectory Prediction Error

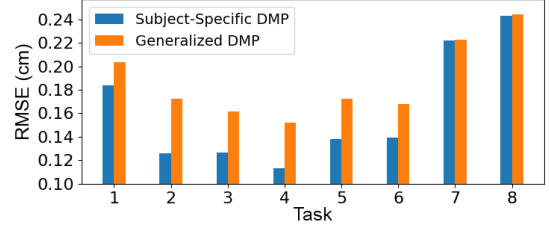


Fig. 5. Mean RMSE computed for both the subject-specific (blue) and generalized (orange) models. On average, subject-specific model produces an RMSE of 0.161 cm while the generalized model produces an RMSE of 0.187 cm (percent difference of 13.90%).

REFERENCES

- [1] M. F. Levin, E. C. Magdalon, S. M. Michaelsen, and A. A. Quevedo, "Quality of grasping and the role of haptics in a 3-d immersive virtual reality environment in individuals with stroke," *IEEE Trans Neural Syst Rehabil Eng*, vol. 23, no. 6, pp. 1047–55, 2015.
- [2] M. Markovic, S. Dosen, C. Cipriani, D. Popovic, and D. Farina, "Stereovision and augmented reality for closed-loop control of grasping in hand prostheses," *J Neural Eng*, vol. 11, no. 4, p. 046001, 2014.
- [3] X. Tan, Y. Li, and Y. Gao, "Combining brain-computer interface with virtual reality: review and prospect," in *IEEE Int Conf Comput Comm*. IEEE, 2017, pp. 514–8.
- [4] L. P. Maletsky, J. Sun, and N. A. Morton, "Accuracy of an optical active-marker system to track the relative motion of rigid bodies," *J Biomech*, vol. 40, no. 3, pp. 682–5, 2007.
- [5] D. Webster and O. Celik, "Systematic review of kinect applications in elderly care and stroke rehabilitation," *J NeuroEng Rehabil*, vol. 11, no. 1, p. 108, 2014.
- [6] N. Johnson, J. Carey, B. Edelman, A. Doud, A. Grande, K. Lakshminarayan, and B. He, "Combined rtms and virtual reality brain-computer interface training for motor recovery after stroke," *J Neural Eng*, vol. 15, no. 1, p. 016009, 2018.
- [7] C. Hunt, R. Yerrabelli, C. Clancy, L. Osborn, R. Kaliki, and N. Thakor, "Pham: prosthetic hand assessment measure," in *Conf Proc Myoelectric Control Symp*, 2017, pp. 221–4.
- [8] M. Masters, L. Osborn, N. Thakor, and A. Soares, "Real-time arm tracking for hmi applications," in *IEEE Int Conf Wearable Implant Body Sensor Networks*. IEEE, 2015, pp. 1–4.
- [9] G. Ligorio and A. M. Sabatini, "A novel kalman filter for human motion tracking with an inertial-based dynamic inclinometer," *IEEE Trans Biomed Eng*, vol. 62, no. 8, pp. 2033–43, 2015.
- [10] M. El-Gohary, L. Holmstrom, J. Huisinga, E. King, J. McNames, and F. Horak, "Upper limb joint angle tracking with inertial sensors," in *Conf Proc IEEE Eng Med Biol Soc*. IEEE, 2011, pp. 5629–32.
- [11] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans Automat Control*, vol. 53, no. 5, pp. 1203–18, 2008.
- [12] S. Raza, P. Misra, Z. He, and T. Voigt, "Building the internet of things with bluetooth smart," *Ad hoc Networks*, vol. 57, pp. 19–31, 2017.
- [13] S. Alavi, D. Arseneault, and A. Whitehead, "Quaternion-based gesture recognition using wireless wearable motion capture sensors," *Sensors*, vol. 16, no. 5, p. 605, 2016.
- [14] J. C. Chou and M. Kamel, "Finding the position and orientation of a sensor on a robot manipulator using quaternions," *Int J Rob Res*, vol. 10, no. 3, pp. 240–54, 1991.
- [15] A. Sharma et al., "A mixed-reality training environment for upper limb prosthesis control," in *IEEE Biomed Circuits Syst Conf*. IEEE, 2018, In Press.
- [16] R. M. Murray, *A Mathematical Introduction to Robotic Manipulation*. CRC press, 2017.
- [17] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, "Evaluation of pose tracking accuracy in the first and second generations of microsoft kinect," in *IEEE Int Conf Healthc Inform*. IEEE, 2015, pp. 380–9.
- [18] G. Hotson, R. J. Smith, A. G. Rouse, M. H. Schieber, N. V. Thakor, and B. A. Wester, "High precision neural decoding of complex movement trajectories using recursive bayesian estimation with dynamic movement primitives," *IEEE Robot Autom Lett*, vol. 1, no. 2, pp. 676–83, 2016.
- [19] M. Kawato, "Internal models for motor control and trajectory planning," *Curr Opin Neurobiol*, vol. 9, no. 6, pp. 718–27, 1999.