Machine Learning for Placement-insensitive Inertial Motion Capture

Xuesu Xiao and Shuayb Zarar

***Abstract*— Although existing inertial motion-capture systems work reasonably well ($\leq 10^{\circ}$ error in Euler angles), their accuracy suffers when sensor positions change relative to the associated body segments ($\pm 60^{\circ}$ mean error and 120° standard deviation). We attribute this performance degradation to undermined calibration values, sensor movement latency and displacement offsets. The latter specifically leads to incongruent rotation matrices in kinematic algorithms that rely on rotational transformations. To overcome these limitations, we propose to employ machine-learning techniques. In particular, we use multi-layer perceptrons to learn sensor-displacement patterns based on 3 hours of motion data collected from 12 test subjects in the lab over 215 trials. Furthermore, to compensate for calibration and latency errors, we directly process sensor data with deep neural networks and estimate the joint angles. Based on these approaches, we demonstrate up to 69% reduction in tracking errors.**

I. Introduction

Since its introduction two decades ago, motion-capture technology has revolutionized a variety of applications in robotics, computer graphics, virtual reality, rehabilitation engineering and athletic training. The increasing popularity of this technology has motivated researchers to keep improving capture accuracy [1]–[4], and experiment with reconstruction techniques such as filtering, optimization and physics-based approaches [5]–[7]. One particular line of research has focused on making motion-capture technology accessible outside of traditional indoor studio-like environments, enabling relatively non-intrusive tracking [8]. Inertial sensing is a technique that makes this approach possible since it eliminates the need for external sensors such as cameras or localization rigs. One limitation of existing inertial-sensing systems, however, is that they require sensors to be precisely positioned on human body parts with the use of either elastic straps or snugly-fitting custom-built clothing [8]–[10]. While allowing motion-capture to take place outside of studios, the necessity to brace sensing devices firmly to the body is still considered intrusive. Furthermore, tight mounting of sensors precludes long-term tracking in a natural way.

Emerging systems have begun the integration of sensors into loose garments, allowing us to overcome the mounting limitations of inertial motion-capture technology [11]–[13]. However, signals in such systems are heavily corrupted with motion artifacts [14]. Thus, tracking accuracies are extremely poor; mean Euler angle errors of up to $\pm 60^{\circ}$ and standard deviations of 120° are observed. In this paper, we use machine-learning algorithms to overcome the challenges posed by noise in garment-integrated inertial motion-capture systems. Fig. 1 shows the performance of our system in comparison with existing approaches of using kinematics-based motion capture. From the figure, and video illustrations that accompany this paper, we observe that our system performance is robust even in the presence of large displacement between IMU sensors and body segments. The following are the specific contributions that we make:

Xuesu Xiao is with the CSE Dept., Texas A&M University, College Station, TX 77843, xiaoxuesu@tamu.edu

Shuayb Zarar is with Microsoft Research, Redmond, WA 98052, shuayb@microsoft.com

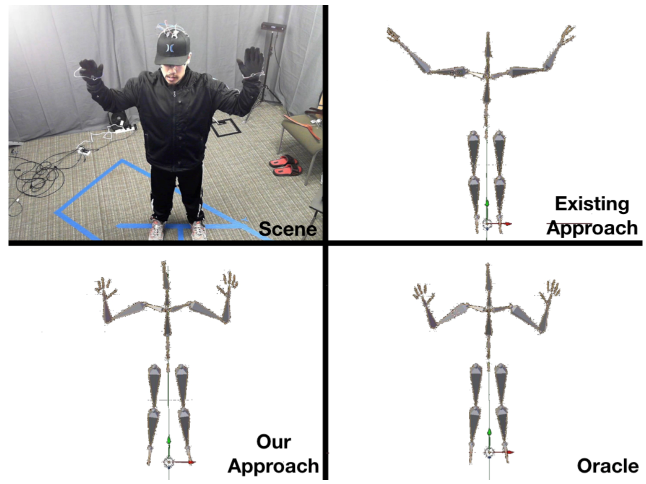


Fig. 1: We tackle issues that arise when sensors move relative to the associated body segments. Such displacements are extreme in case of garment-integrated sensors.

- We demonstrate different methods of combining inertial and infrared (IR) proximity sensors for motion capture. IR sensors provide stable measures of distance, while inertial sensors suffer from drift.
- We fuse sensor data within a deep neural network (DNN) model that learns displacement patterns between inertial measurement units (IMUs) and body segments when there is relative motion between them.
- We propose a machine-learning model that trades generalizability for accuracy. It directly exploits raw sensor data to overcome latency and calibration issues.

The rest of the paper is organized as follows. In Sec. II, we provide an overview of the state-of-the-art in inertial motion-capture systems along with background on kinematics-based pose tracking. In Sec. III, we describe our proposed machine-learning techniques of modeling sensor-displacement patterns within the kinematics algorithm and directly estimating joint-angles with neural networks. In Sec. IV, we present tracking results obtained from the proposed methods and compare them with kinematics-based estimation. In the same section, we discuss advantages and disadvantages of the proposed methods and identify directions for future research. Finally, we conclude in Sec. V.

II. Background and Related Work

In this section, we provide an overview of existing techniques for motion capture. We also present a kinematics-based algorithm that we utilize in this paper.

A. *Limitations of Inertial Motion Capture*

Traditional camera-based motion-capture systems constrain tracking to a small capture volume and are sensitive to occlusions. Although multiple cameras alleviate the occlusion problem, they incur substantial complexity and are limited in mobility [15]. Inertial motion-capture systems have aimed to address the occlusion problem by employing techniques of sensor fusion and numerical optimization [5], [6], [16]. Further, to tackle the mobility issue, sparse collections of sensors have been utilized within custom-built compression suits [8]–[10]. Although, combinations of such systems permit occlusion-free, far-field pose tracking, they are highly intrusive. Extra sensors still need to be firmly affixed to the body making it impractical for long-term tracking. Emerging systems have proposed to handle this issue by embedding sensors into everyday loose-fitting clothes [11], [12]. This provides a natural way to merge motion-capture technology into our lives. However, these approaches introduce movement artifacts that can hurt the accuracy of even the most sophisticated tracking algorithms.

To the best of our knowledge, there is little existing work that improves accuracy of inertial motion capture in the presence of sensor noise. Furthermore, there are only a limited number of efforts that employ machine-learning techniques to improve tracking accuracy. For good performance during complex motions, researchers have explored dynamic shifting of root points from pelvis to the feet [1]. Support-vector regression has been used to make estimation of knee angles more accurate during walking [2]. Including dynamics and combining optical and inertial sensors has also helped increase tracking accuracy [3], [4], [7]. In this paper, we propose to employ deep-learning techniques to improve tracking accuracy in the presence of signal noise (due to motion artifacts). Before we present our approach, we describe a kinematics-based algorithm that we utilize.

B. *Kinematics-based Pose Tracking*

Kinematics-driven tracking is a known approach for inertial motion capture [12], [17], [18]. It depends on a rotational transformation of points in 3D space. For instance, any point can be transformed from frame X to frame Y *via* the rotation matrix R^Y_X. Suppose B_i, S_i and G represent coordinate frames associated with the body segment i, corresponding sensor on the body segment i, and the global reference (aligned with the Earth's magnetic field) at calibration time, respectively. Further suppose B'_i and S'_i represent the body and sensor frames after arbitrary motion. We can express any imaginary point P *via* a transformation of coordinate systems between body segments i and j as follows:

$$R^G_{S'_i} R^{S'_i}_{B'_i} R^{B'_i}_{B'_j} P = R^G_{S'_j} R^{S'_j}_{B'_j} P \tag{1}$$

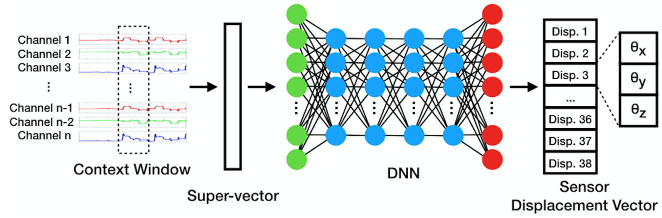


Fig. 2: DNN architecture for modeling sensor displacement patterns during limb movement.

By simplifying this expression with matrix inversion, we get:

$$R^{B'_i}_{B'_j} = R^{B'_i}_{S'_i} {R^G_{S'_i}}^{-1} R^G_{S'_j} {R^{B'_j}_{S'_j}}^{-1} \tag{2}$$

$R^G_{S'_i}$ and $R^G_{S'_j}$ are the outcomes of the IMU sensor-fusion algorithm computed dynamically [19]. $R^{B'_i}_{S'_i}$ and $R^{B'_j}_{S'_j}$ represent displacement between sensors and body parts. They are obtained during calibration and assumed to remain the same through the course of motion. This assumption is valid if sensors are rigidly mounted to the body. Thus, by measuring $R^G_{S'_i}$ and $R^G_{S'_j}$, we can compute the joint angle $R^{B'_i}_{B'_j}$ between body parts i and j.

III. Machine-learning to Improve Tracking Performance

In this section, we propose techniques to quantify sensor displacements on the body and utilize them within the kinematics algorithm for pose tracking. We also explore direct joint-angle estimation with neural networks.

A. *Learning Sensor-displacement Patterns*

The main issue when sensors move relative to body segments is that the rotation matrices $R^{B'_i}_{S'_i}$ and $R^{B'_j}_{S'_j}$ in Eq. (2) cannot be assumed to equal $R^{B_i}_{S_i}$ and $R^{B_j}_{S_j}$, which are measured at calibration time. These matrices represent transformations between the sensor and body segment coordinate frames. If mounting is not rigid, they change depending on how sensors are displaced during the course of motion. Our strategy is to use data-driven techniques to model this displacement behavior.

Suppose, we have knowledge of joint angles through other means (such as an external optical tracker), we can re-arrange terms in Eq. (2) to obtain the following relationship:

$$R^{B'_j}_{S'_j} = {R^{B'_i}_{B'_j}}^{-1} R^{B'_i}_{S'_i} {R^G_{S'_i}}^{-1} R^G_{S'_j} \tag{3}$$

We fix the kinematic root on the *head* and assume that there is minimal displacement offset of the sensor with respect to the body segment at this point. Thus, for this specific body segment, we assume $R^{B'_i}_{S'_i}$ to be approximately equal to $R^{B_i}_{S_i}$, which is the transformation matrix obtained at the time of calibration. We proceed to solve Eq. (3) for $R^{B'_j}_{S'_j}$, which is the body segment connected to the head *i.e.*, middle spine. Similarly, we continue along the kinematic chain up to 4 body extremities and estimate values of the displacement matrices $R_{S'_j}^{B'_j}$ at different limb positions. These form the ground-truth labels in our modeling framework.

We formulate a regression setup based on the deep neural-network architecture shown in Fig. 2. For the network, we build an input vector that comprises raw readings from one IR proximity sensor and 10 IMU data channels per body part (3 values each from the accelerometer, gyroscope and magnetometer and 1 value from the temperature sensor). These vectors are formed at a fast frame rate (typically 30 per second). Further, to account for temporal context, we concatenate additional such vectors that appear a few steps before the current time leading to an extended super-vector of dimensionality $11 \times m \times n$, where m is the number of body segments being tracked and n is the number of additional context windows being used. The output of the network is a set of three Euler angles corresponding to each body segment. Thus, the output of our network is a vector $\hat{y}$ with dimensionality $3 \times m$. To set up the regression cost function, we determine the sensor-displacement matrices $R_{S'_j}^{B'_j}$ by solving Eq. (3) and transform them to Euler angles, yielding a target vector y that also has dimensionality $3 \times m$. Thus, we train the neural network to minimize the following mean absolute error (MAE) function:

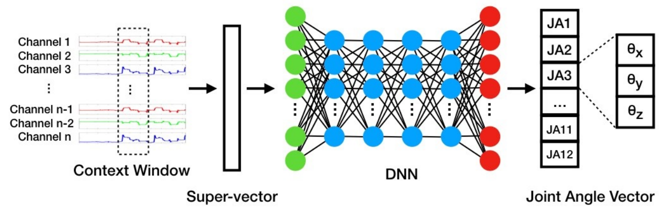
$$\underset{\hat{y}_i}{\operatorname{argmin}} \frac{1}{3m} \sum_{i=1}^{3m} |\hat{y}_i - y_i| \tag{4}$$

Once the model is trained, we apply it to the test data at inference time to predict sensor displacements $R_{S'_1}^{B'_1}, R_{S'_2}^{B'_2}, \dots R_{S'_m}^{B'_m}$ (obtained after transforming Euler angles to rotation matrices). We plug these displacement values into Eq. (2) to compute the joint angles $R_{B'_j}^{B'_i}$. Thus, our system naturally fuses information from the IMU and IR sensors to account for sensor displacement during limb motion.

B. End-to-End Joint Angle Estimation

One limitation of modeling sensor displacement is that it does not account for other sources of error in the inertial-tracking system such as sensor-movement latency and calibration inaccuracies. To overcome these challenges, we propose to exploit the abstraction power of neural-networks by directly estimating joint-angles without the kinematic constraints.

The modified DNN architecture is shown in Fig. 3. It utilizes the same input super-vector as the previous model. However, the output vector is the set of joint angles $R_{B'_j}^{B'_i},\ i,j \in [1,m]$ and $|i-j| = 1$, instead of the sensor displacement values $R_{S'_1}^{B'_1}, R_{S'_2}^{B'_2}, \dots R_{S'_m}^{B'_m}$ as in the previous model. Thus, the output vector has dimensionality $3 \times (m-1)$. The ground-truth labels for the joint angles are still obtained from an optical tracking system. The cost function in this case is the MAE computed between the Euler values of the estimated and ground-truth joint angles.



Layer No.	Layer Type	IO Shape	No. Parameters
1	Dense	2660, 128	340608
2	Dense	128, 128	16512
3	Dense	128, 128	16512
4	Dense	128, 36	4644
		Total	378276

Fig. 3: DNN architecture and parameter values for directly estimating joint angles from raw sensor data.

IV. Experimental Results

In this section, we study the static and dynamic performance of the two machine-learning models in comparison with the kinematics-based algorithm, which incorrectly assumes that sensors are not moving with respect to the body segments. We also explore optimization of network parameters.

Data and implementation. We collected data with the use of a set of clothes (jacket, pants, hat, gloves and shoes) that had IMU and IR sensors integrated into them. Data from these sensors was streamed over a WiFi user datagram protocol (UDP) network, filtered, interpolated and re-sampled to 30 samples-per-second to simultaneously match the frame rate of our optical tracking system. We fused data from two Microsoft Kinect sensors *via* a third-party software to obtain the ground-truth optical joint-angle estimates [20]. We tracked 12 joint angles ($m = 13$ body segments) using a total of 143 sensor channels (11×13 body segments) and 4290 sensor samples-per-second (143×30). After an internal approval process, we invited 12 adult subjects (8 male and 4 female) to perform 5 motion patterns in a calibrated studio for a total duration of 3.5 hours. The motion patterns comprised upper and lower arm lifts to reach multiple extremities, arm swings and periodic movements such as boxing and walking. We avoided motion along the axis of bones since Kinect is not capable of tracking this type of motion. We also collected data samples for arbitrary motion patterns in order to study the generalization capacity of the neural-network models. Overall, we collected over half a million data samples *via* 215 trials.

We used Theano with Keras APIs for the neural networks. The model parameters are shown in Fig. 3. We used Xavier normal initializer for the neurons and rectified linear unit (ReLU) activation function after each dense layer. We implemented the proposed models on a PC with 64 GB DDR4 RAM, 2.8 GHz, 16 core 2× Intel Xeon CPUs, and an Nvidia Titan X (Pascal) GPU with 12 GB DDR5 RAM and 3584 Cuda cores running at 1.5 GHz. We utilized 64, 16 and 20% of the collected data as training, validation and test sets, respectively. To avoid over-fitting, we tracked loss profiles of both the training and validation sets. An example of these scores for one of our network models is shown in Fig. 4.

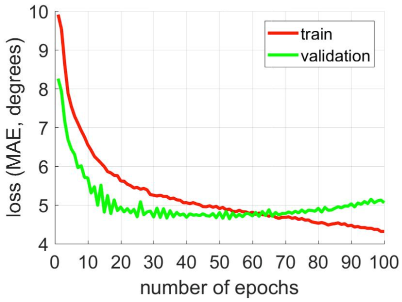


Fig. 4: Consistent trends in training and validation losses indicate no over-fitting of our models.

From the figure, we see that the validation loss follows the training loss ensuring model resilience to over-fitting.

A. *Hyper-parameter Tuning*

We perform sparse grid search to find the best parameters for our models. Due to space constraints, we only discuss parameter optimization for the second method of end-to-end joint-angle estimation. We observe that there are 3 parameters that impact network performance: number of dense layers, number of neurons per layer, and size of the context window. We experimented with these along with a selection of different sensor modalities (more than 1 IMU per body segment). Eventually, the optimal DNN architecture comprised 4 dense layers, 128 neurons per layer and a 7-frame context window. The best results also came from utilizing more than one IMU per body segment, while leaving out the IR sensors. To ease visualization of the design space, we fix one parameter to the optimal value and illustrate the behavior of the remaining two. Fig. 5 shows the change in performance with respect to the number of layers, neurons and context-window size. We observe that deeper DNNs can help improve the test score. Furthermore, a larger context window makes the DNN depth more effective.

The impact of sensor modalities for estimation is shown in Fig. 6. We performed the same sweep of network parameters

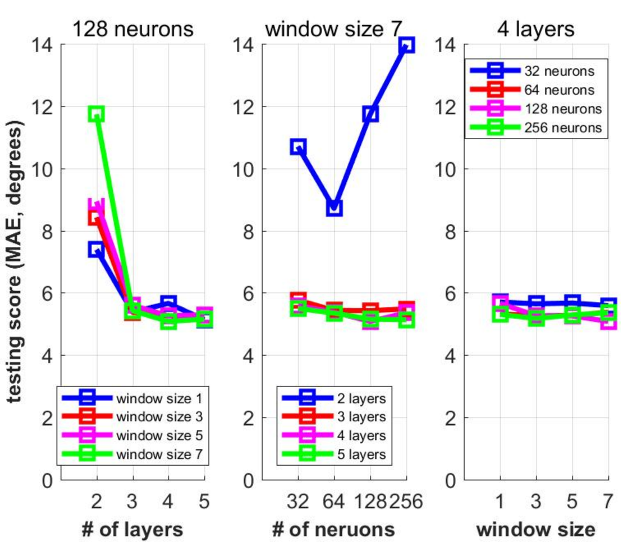


Fig. 5: Network performance with different parameters.

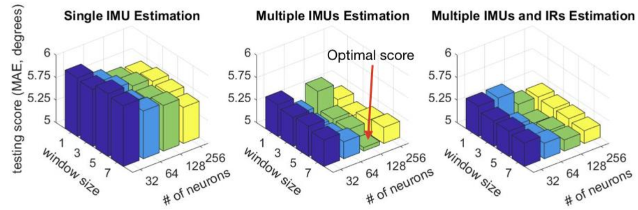


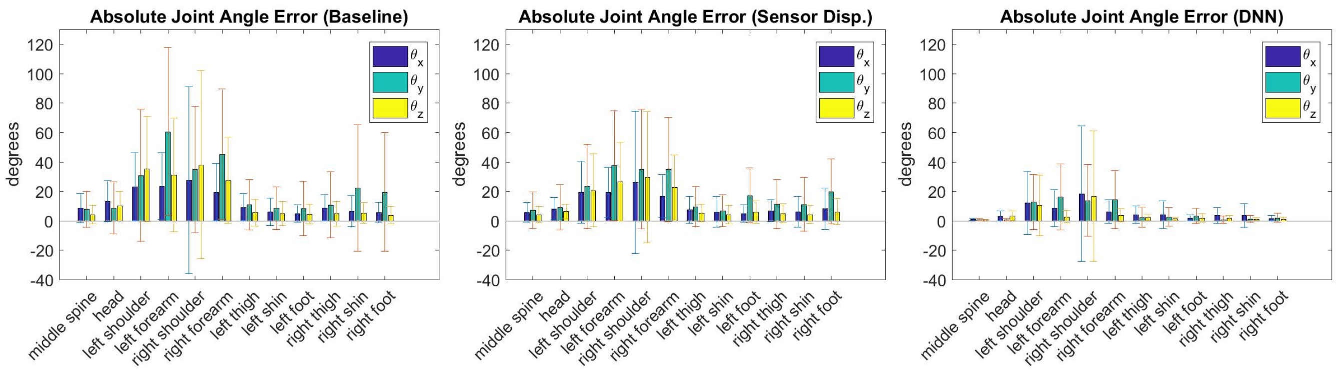
Fig. 6: Network performance improves with extra sensors.

as before while utilizing data from different sensors. We use 1 IMU, 2 IMUs and 2 IMUs + 1 IR sensor per body segment for the plots on the left, center and right of Fig. 6, respectively. We observe that sensor redundancy substantially reduces the tracking errors. Thus, utilizing more than one sensor is an effective approach to mitigate artifacts. An interesting future direction of research would be to determine strategies that would effectively exploit multiple sensor modalities for pose tracking.

B. *Tracking Accuracy*

Fig. 7 shows the absolute joint-angle error from the following three approaches: (a) kinematics-based modeling with the assumption of no sensor displacement, also referred to as the *baseline algorithm*, (b) neural-network based estimation of sensor displacement combined with kinematic modeling, and (c) direct estimation of joint angles with a deep neural network. Although all of them exhibit significant error when tracking upper-limb movements, compensating for sensor displacement helps improve the accuracy of kinematics-based modeling. Furthermore, end-to-end estimation achieves the best precision; tracking accuracy is improved by 69% compared to kinematics-based tracking. Besides MAE, the standard deviation also improves for the proposed approaches. As expected, sensors mounted on a single piece of garment (like the jacket and pants) exhibit error patterns that are highly correlated.

Time-domain performance. Fig. 8 shows the tracking profile of the right shoulder joint angle with the three approaches for one random trial. Three repetitions of the same motion were performed by one subject over 1.4 min. The worst performing baseline approach (red line) suffers from different error sources like invalid calibration, sensor displacement and movement latency (sensor position changes slower than the body segment position during motion). From the figure, we observe that learning sensor displacement (gray line) helps improve tracking accuracy. However, it still suffers from calibration errors (frames 300-800). This is because the neural-network model does not have ground-truth information to find a mapping between sensor data and initial sensor displacement during calibration (T-pose). Our approach helps mitigate errors due to movement latency and sensor displacement. It is worth noting that this approach relies on the output of IMU sensor fusion to compute $R^G_{S'_i}$ and $R^G_{S'_j}$ and estimate the displacement ground-truth labels [Eq. (3)]. Thus, noise from the sensor-fusion algorithm propagates over to the DNN labels, which leads the neural-



(a) Kinematics-based modeling (b) Learning sensor displacement (c) End-to-end joint-angle estimation

Fig. 7: The proposed machine-learning approaches achieve the best tracking accuracy. Compared to kinematics-based tracking, end-to-end joint angle estimation results in a 69% lower MAE in the estimated Euler angles.

network to learn data patterns that are imprecise. As a result, very aggressive spatial filtering is applied to maintain a smooth tracking trajectory. This is apparent from the rising slope observed during the first 300 frames in Fig. 8. The best performance is achieved by the direct joint-angle estimation approach (green line). It completely eliminates the calibration and sensor-displacement errors. Since the estimated pose is accurate during calibration phase, the tracker responds as soon as the ground-truth pose changes. Thus, errors due to sensor movement latency are also mitigated. On motion extremities, there is still some tracking error. However, this approach is still the closest to the ground-truth.

Dynamic response. As expected, the performance of inertial motion-capture is hurt during fast movements. The accuracy degradation is extreme in case of the baseline kinematics algorithm. This is because during fast movements, errors due to sensor displacement and movement latency are amplified. Fig. 9 shows the distribution of tracking errors at different speeds of motion. We compute these results based on a spectral-domain analysis of the joint-angle profiles. From the figure, we observe that our DNN-based direct joint-angle estimator maintains tracking accuracy even during fast movements, which potentially makes it the algorithm of choice for tracking pose in most practical scenarios.

C. Trading Accuracy for Generalizability

Both machine-learning algorithms are capable of generating better tracking accuracy than kinematics-based modeling that is oblivious to sensor displacement. On the one hand, due to the fact that noisy IMU sensor-fusion information is utilized to compute ground-truth rotation matrices, training labels are corrupted with fusion noise. This deteriorates the final tracking accuracy of the DNN that models sensor displacement. Aggressive filtering helps but makes the tracker less responsive. On the other hand, directly estimating joint angles with sensor data can lower average Euler-angle errors to 6°. These gains come at the cost of model generalizability. To study the robustness of the neural-network approaches, we utilized data from subjects and motions that were not part of the collected dataset. Baseline kinematics-based modeling achieves an average root mean-squared error (RMSE) of 20° over all joint angles. The results for the proposed machine-learning approaches are shown in Table. I.

When presented with motions that are not part of the training set, directly estimating joint angles with the neural

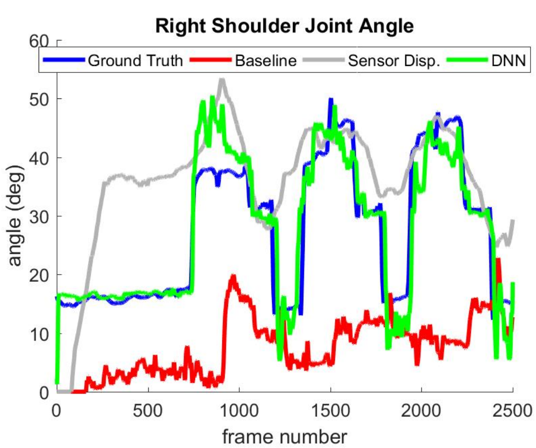


Fig. 8: DNN-based modeling closely tracks the ground-truth pose. End-to-end estimation achieves the best result.

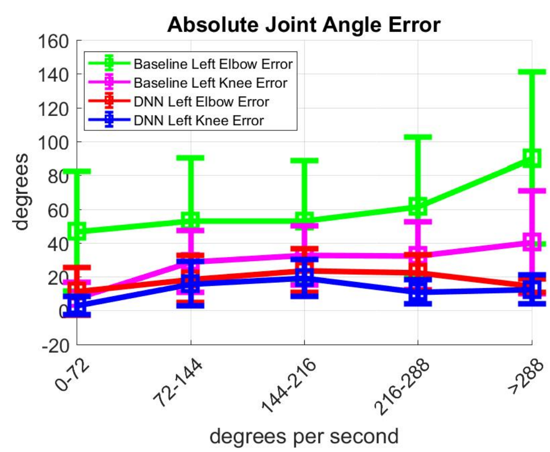


Fig. 9: The DNN models maintain performance even for fast limb movements, while the baseline accuracy is poor.

TABLE I. Average RMSE of joint angles on unseen test data.

	Seen	Unseen
DNN to learn sensor displacement	16.42°	14.54°
DNN to directly estimate joint angles	6.14°	10.17°

network is less accurate. However, this is not the case with the approach of modeling sensor displacement. In fact, in this case the performance is slightly increased; error reduces from $16.42°$ to $14.54°$. We hypothesize that this performance difference is due to the nature and amount of training data that we use. Given the finite set of motion and subject samples, our training set is only able to encode a subset of variance in the large space of possible Euler angles. Thus, a completely data-driven approach, such as the direct joint-angle estimator, suffers from low accuracy on unseen test data. This exposes the limited generalizability of such a method. In contrast, utilizing machine learning to model sensor displacement within a kinematics framework is a hybrid approach that relies on patterns that have lower variance; the space of sensor displacements is smaller than that of Euler angles. Thus, our training dataset potentially has more coverage over this domain. Therefore, unless there is access to sufficient labeled training data, the approach of directly estimating joint-angles may be more suitable for structured tasks such as recognizing specific activities rather than tracking pose. On the other hand, the hybrid approach of modeling sensor displacement may be a good candidate for motion capture in general.

In this paper, we implemented simple DNN models for inertial motion capture. More advanced neural-network architectures could help boost the accuracy further. We utilized temporal information in the data through a simple context window. Recursive neural networks (RNNs) may be effective alternatives to exploiting time-domain relationships. This is especially interesting since motion-capture can be modeled as a sequence-prediction problem with kinematic constraints. To develop these ideas further, there is a dire need for labeled datasets that not only capture motion variations to a comprehensive degree but are also rich in sensor modalities and subject diversity.

V. Conclusions

We developed two machine-learning methods based on deep-neural networks to address the issue of significant motion artifacts during inertial motion capture. These artifacts arise naturally in systems where sensors are allowed to move freely without constraints on mounting, such as in the case of sensors integrated into garments. We demonstrated superior tracking accuracy of both approaches when compared to a kinematics-based algorithm that is oblivious to sensor displacements during motion capture. Specifically, we showed that direct estimation of joint angles with DNNs can lower RMSE from $20°$ to $6°$. Based on a study of the dynamic response of the algorithms, we found that these low errors are maintained even in the presence of drastic motion. We identified that a large amount of motion data is needed to achieve good generalizability of the DNN models. However, the hybrid approach of learning sensor-displacement patterns within a kinematics algorithm can be a viable alternative when we lack large volumes of labeled training data.

References

[1] Y. Zheng, K.-C. Chan, and C. C. Wang, "Pedalvatar: An IMU-based real-time body motion capture system using foot rooted kinematic model," in *Int. Conf. Intelligent Robots and Systems*. IEEE, Sep. 2014, pp. 4130–4135.

[2] S. Ahuja, W. Jirattigalachote, and A. Tosborvorn, "Improving accuracy of inertial measurement units using support vector regression," *Stanford University, Tech. Rep*, Oct. 2016.

[3] E. Foxlin, M. Harrington, and G. Pfeifer, "Constellation: A wide-range wireless motion-tracking system for augmented reality and virtual set applications," in *Proc. Annual Conf. Computer Graphics and Interactive Techniques*. ACM, Jul. 1998, pp. 371–378.

[4] T. von Marcard, G. Pons-Moll, and B. Rosenhahn, "Human pose estimation from video and IMUs," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 38, no. 8, pp. 1533–1547, Aug. 2016.

[5] R. Zhu and Z. Zhou, "A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package," *IEEE Trans. Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. 295–302, Jun. 2004.

[6] M. Kok, J. D. Hol, and T. B. Schön, "An optimization-based approach to human body motion capture using inertial sensors," *Proc. Int. Federation of Automatic Control*, vol. 47, no. 3, pp. 79–85, Aug. 2014.

[7] S. Andrews, I. H. Casado, T. Komura, L. Sigal, and K. Mitchell, "Real-time physics-based motion capture with sparse sensors." in *European Conf. Visual Media Production*, Dec. 2016, pp. 1–5.

[8] D. Vlasic *et. al.*, "Practical motion capture in everyday surroundings," in *ACM Trans. Graphics*, vol. 26, no. 3. ACM, Jul. 2007, p. 35.

[9] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors," *Xsens Motion Technologies, Tech. Rep*, Jan. 2009.

[10] T. von Marcard, B. Rosenhahn, M. J. Black, and G. Pons-Moll, "Sparse inertial poser: Automatic 3D human pose estimation from sparse IMUs," in *J. Computer Graphics Forum*, vol. 36, no. 2. Wiley Online Library, May 2017, pp. 349–360.

[11] H. Harms, O. Amft, and G. Troster, "Influence of a loose-fitting sensing garment on posture recognition in rehabilitation," in *Int. Conf. Biomedical Circuits and Systems*. IEEE, Nov. 2008, pp. 353–356.

[12] X. Xiao and S. Zarar, "A wearable system for articulated human pose tracking under uncertainty of sensor placement," in *under review*, 2018.

[13] C. Mattmann, O. Amft, H. Harms, G. Troster, and F. Clemens, "Recognizing upper body postures using textile strain sensors," in *Int. Symp. Wearable Computers*. IEEE, Oct. 2007, pp. 29–36.

[14] S. Mohammed and I. Tashev, "Unsupervised deep representation learning to remove motion artifacts in free-mode body sensor networks," in *Int. Conf. Body Sensor Networks*. IEEE, May 2017, pp. 183–188.

[15] M. Windolf, N. Götzen, and M. Morlock, "Systematic accuracy and precision analysis of video motion capturing systems-exemplified on the Vicon-460 system," *J. Biomechanics*, vol. 41, no. 12, pp. 2776–2780, Jan. 2008.

[16] D. H. Kang, Y. J. Jung, A. J. Park, and J. W. Kim, "Human body motion capture system using magnetic and inertial sensor modules," in *Proc. Int. Universal Communication Symp.*, Oct. 2011, pp. 1–6.

[17] J. Favre, R. Aissaoui, B. M. Jolles, J. A. de Guise, and K. Aminian, "Functional calibration procedure for 3D knee joint angle description using inertial sensors," *J. biomechanics*, vol. 42, no. 14, pp. 2330–2335, Oct. 2009.

[18] G. Pons-Moll, A. Baak, J. Gall, L. Leal-Taixe, M. Mueller, H.-P. Seidel, and B. Rosenhahn, "Outdoor human motion capture using inverse kinematics and von mises-fisher sampling," in *Int. Conf. Computer Vision*. IEEE, Nov. 2011, pp. 1243–1250.

[19] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Int. Conf. Rehabilitation Robotics*. IEEE, Jul. 2011, pp. 1–7.

[20] iPi Soft LLC. iPi Soft motion capture for the masses. [Online]. Available: http://ipisoft.com/