Meta-Pose: Environment-adaptive Human Skeleton Tracking with RFID

†Chao Yang, †Lingxiao Wang, ‡Xuyu Wang, and †Shiwen Mao
†Dept. of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201
‡Dept. of Computer Science, California State University, Sacramento, CA 95819-6021
Email: czy0017@auburn.edu, lzw0039@auburn.edu, xuyu.wang@csus.edu, smao@ieee.org

***Abstract*—Human pose tracking has attracted great interest recently. Considerable efforts have been made in Radio-Frequency (RF) sensing techniques for human pose tracking without using a video camera. Although the existing RF based schemes can well protect user privacy, they are usually sensitive to the RF environment and are hard to generalize to new environments. In this paper, we analyze the challenges of generalization of Radio-Frequency Identification (RFID) based human pose tracking systems. We then present an RFID based 3D human pose tracking system, termed *Meta-Pose*, which incorporates meta-learning and few-shot fine-tuning to achieve high adaptability to new environments. The proposed system is implemented with commodity RFID devices and extensive experiments are conducted for performance evaluation. The experiment results validate the superior human pose tracking performance and high adaptability of the proposed Meta-Pose system.**

***Index Terms*—3D human pose tracking, few-shot fine-tuning, generalization, meta-learning, RFID sensing.**

I. Introduction

Human pose tracking has attracted great interest in recent years, because it is highly useful for numerous applications such as human-computer interaction, video surveillance, and somatosensory games. The advances in human pose tracking have been driven by the new developments in computer vision, from two-dimensional (2D) poses [1] to the three-dimensional (3D) realtime pose tracking [2]. However, such vision based schemes often raise security and privacy concerns. For example, it has been reported that millions of wireless security cameras were possibly hacked [3]. The collected video data for pose tracking could be illegally intercepted. Several radio frequency (RF) sensing schemes have been proposed to address the privacy concern in human pose tracking [4], including Frequency-Modulated Continuous Wave (FMCW) radar based [5], mmWave radar based [6], WiFi-based [7], [8], and RFID-based schemes [9]–[11]. Compared with vision based techniques, RF sensing based pose tracking has no requirement for the lighting condition, and the privacy of users can be well protected.

In RF based pose generation systems, deep learning techniques are usually used to transform sampled RF data to human pose. However, such machine learning based techniques usually have the generalization problem when applying a well-trained model in a new, unknown environment. Since RF signals propagate in the open air, the collected RF data are sensitive to the changes in the environment, such as the antenna deployment, the layout and obstacles of the surroundings, and moving objects/subjects nearby. Under such environment changes, the same human subject could generate considerably different RF features when tested in different environments. Developing human pose estimation techniques that are adaptive to the environment has become a great challenge for RF based techniques.

When applying the well-trained (or, pretrained) model to a new, unknown environment, we can fine-tune the model by further training it with new data collected from the unknown environment, such that the specific features of the new domain can be better captured. For good generalization performance, the amount of data used for fine-tuning should be as low as possible, in order to minimize the time, effort, and cost of obtaining training data from the new environment. This is important for the system to be easily deployed in practice. To this end, meta-learning, a.k.a. "learning to learn" [12] provides an excellent solution. Meta-learning optimizes the neural network based on different learning tasks or datasets [13], so the network will be appropriately initialized and be amenable for adaptation to new environments. When transferred to a new RF environment, the meta-learning model will only require a few training examples from the new environment for fine-tuning (i.e., few-shot fine-tuning).

In this paper, we tackle the environment adaptation challenge with a meta-learning approach and propose a novel environment-adaptive, RFID based 3D human skeleton tracking system termed *Meta-Pose*. As prior work RFID-Pose [10], the system leverages RFID tags attached to the human body to capture the movements of human body parts. It is also a vision-assisted scheme, where Kinect generated vision data is used for supervised training. However, vision data will not be needed for inference, so there will be no privacy concerns. To address the generalization problem, we first analyze the main causes for the divergence of RFID data in different RF environments. Based on the analysis, we then propose a novel Meta-Pose initialization algorithm to pretrain the model with RFID data sampled from a few environments. With few-shot fine-tuning, the Meta-Pose system is able to accurately track 3D human skeleton in a new, unknown environment. Extensive experiments are conducted to validate the high environment adaptation ability of the proposed Meta-Pose system.

The main contributions of this paper are summarized in the following.

- To the best of our knowledge, Meta-Pose is the first

978-1-7281-8104-2/21/$31.00 ©2021 IEEE

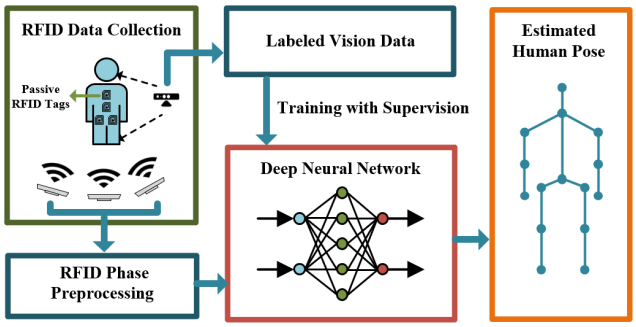

Fig. 1. Overview of the proposed RFID pose tracking system.

environment-adaptive 3D human pose estimation system designed with commodity RFID reader and tags, which can be easily deployed to track 3D human skeletons with RFID data in any RF environment.

- We analyze the divergence of RFID data in different environments and identify the main challenges to the generalization of RFID based techniques, including sensitivity divergence of RFID tags and phase distortion for different sampling environments.
- We propose a novel Meta-Pose initialization algorithm to pretrain the deep learning model with data sampled from a small number of known environments. The algorithm is based on the meta-learning framework, and a domain fusion technique is incorporated to generate more fake environments to better train the model. The pretrained network can be quickly adapted to new data sampled from a new environment.
- We develop a prototype system with commodity RFID tags/reader, where Kinect 2.0 is used to obtain ground truth data for training the model. The performance of Meta-Pose is evaluated with extensive experiments as well as comparison with a baseline scheme [10]. The experimental results demonstrate that the proposed Meta-Pose system can accurately track 3D human skeletons with high environmental adaptability.

In the remainder of this paper, an overview of the proposed system is presented in Section II. Section III examines the challenges of the generalization problem. Section IV presents the Meta-Pose solution to the challenges. Our prototype implementation and experimental study are presented in Section V. Section VI summaries this paper.

II. Overview of the Proposed System

The Meta-Pose system is proposed to estimate 3D human pose with RFID data collected from the passive RFID tags attached to the human subject. An overview of the Meta-Pose system is shown in Fig. 1. The system is composed of three key components, including (i) RFID data collection, (ii) RFID phase preprocessing, and (iii) a deep neural network.

A. Phase Data Collection and Preprocessing

In the RFID pose tracking system, human pose is learned from RFID phase data, which is obtained by interrogating the tags with the RFID Low Level Reader Protocol (LLRP). Since the RF signal is sent from the antenna, reflected by the passive RFID tag, and received by the antenna, the received RFID phase value Θ is given by [14]:

$$\Theta = \frac{2\pi 2 R f_c}{v} + \Theta_c,\ c = 1, 2, ..., 50, \tag{1}$$

where R is the distance of the LOS path between the reader antenna and tag, and c is the channel index, which changes from 1 to 50 every 200ms in Ultra High Frequency (UHF) RFID systems following the FCC regulation [14].

Next, the RFID phase data should be prepossessed to mitigate the impact of the random Θ_c on different channels. To this end, the phase variation Φ between two adjacent samples would be effective, which is calculated as:

$$\begin{aligned}\Phi(n) &= \Theta(n) - \Theta(n-1) \\ &= \frac{2\pi 2(R(n) - R(n-1))f_c}{v}, c = 1, 2, ..., 50, n > 1,\end{aligned} \tag{2}$$

where n is the sample index on each channel and $R(n)$ is the propagation distance corresponding to the nth sample on channel c. As (2) shows, the impact of the random channel hopping offset Θ_c has been effectively removed from the phase variation Φ. The phase variation only depends on the distance of the LOS propagation path $R(n)$. Therefore, the sequence of phase variations $\{\Phi_2, \Phi_3, ...,\}$ can be translated into a sequence of antenna-tag distances $\{R_2, R_3, ...,\}$, which records the trajectory of human body movements. Consequently, with multiple tags attached on the human body, the RFID phase variations for the attached tags can be leveraged to construct the human skeleton and track 3D human poses.

B. Deep Neural Network for Pose Training

Although phase variation can effectively capture the movements of the tags attached to human body, the translation from phase variation data to 3D human pose is still a challenge. In several existing RFID based human pose tracking systems, the transformation is mostly accomplished with deep learning techniques [10], [11], which is mainly composed of a recurrent autoencoder and a forward kinematic layer. The brief structure of the deep learning model is presented in Fig. 2. As the figure shows, the network is designed to generate a 3D human pose sequence, consisting of coordinates data, from received RFID data. The recurrent encoder is to extract both long-term and short-term features from the RFID data sequence, which is used as input to the following recurrent decoder. With a given initial skeleton, the decoder layer will transfer the features of the RFID data sequence to a quaternion sequence.

Rather than using RF signals to generate a confidence map for human skeleton reconstruction [1], [7], RFID based pose tracking system is designed to estimate human pose with the forward kinematic technique, which is widely used in the robotics and 3D animation [15]. This is because the information rate of the RFID system is too low to generate a confidence map with an acceptable resolution. The forward kinematic technique, however, only requires the quaternions for the joints of the human skeleton.

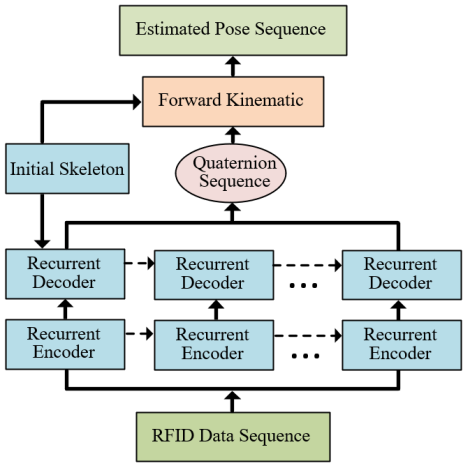

Fig. 2. Structure of the deep learning model used in RFID based 3D human pose tracking.

As RFID-Pose [10] and Cycle-Pose [11], vision data collected by a Kinect 2.0 are used as labels for supervised training. The network is trained with a loss function that computes the difference between the estimated pose and the labeled vision data sampled simultaneously when the RFID data is collected, so the well-trained network can effectively transform RFID data sequence to 3D pose sequence [10].

III. Generalization Challenges

To analyze the influence from the environment, the RF data sampled from a different environment is considered to be from a different data domain. In RF sensing systems, the sampling environment depends on the characteristics of all the propagation paths. However, as given in (1), the received RFID signal is mainly determined by the LOS path due to its near-field communications nature. The interference of the surroundings is thus limited in RFID systems, and the data domain of RFID sensing systems is due to two main causes: (i) tag sensitivity divergence and (ii) phase data distortion.

A. Sensitivity Divergence in Different Data Domains

The first cause of data divergence in different data domains is the variation of tag sensitivity. When multiple tags are scanned by one antenna, some tags are more likely to be detected, while other tags may hardly be scanned by the antenna. We define *tag sensitivity* as the possibility of being successfully detected by the antenna, which mainly depends on the received power strength of each tag.

Following the Friis transmission formula, the received power S_r from a passive RFID tag can be represented by [16]:

$$S_r = G_{An} G_{Tag} L \left(\frac{\lambda_c}{4\pi R} \right)^4 S_t, \tag{3}$$

where S_t is the reader's transmit power; G_{An} and G_{Tag} are the power gains of the transmitter antenna and the tag, respectively; L represents the aggregated attenuation coefficient, accounting for losses at the antenna cable and polarization, etc. during the transmission process; λ_c is the wavelength of the current channel c, and R is the LOS path distance as mentioned in (1). Eq. (3) shows that with the same antenna and same tag type, G_{An} and G_{Tag} could be considered as a constant, so the received power strength is mainly degraded by an increased LOS path distance R and the attenuation loss L.

When applying a trained deep learning model to a different data domain, the inference performance could be poor, since the tag sensitivity in the new data domain could be very different from where the model was trained.

B. Phase Distortion in Different Data Domains

The second cause for data domain divergence is the phase distortion caused by different antenna deployment scenarios. As (1) shows, the phase data of each tag is determined by the LOS propagation path distance R, which is the length of the space vector $\vec{R}$. For the tags attached to a moving human body, we can consider the overall space vector as the sum of two subspace vectors as: $\vec{R} = \vec{R}_s + \vec{R}_d$, where $\vec{R}_s$ is the *static vector* determined by the deployment scenario and the *dynamic vector* $\vec{R}_d$ is generated by the movements of the subject. According to (1), the sampled phase Θ is affected by both $\vec{R}_s$ and $\vec{R}_d$ as:

$$\Theta = \frac{2\pi 2|\vec{R}_s + \vec{R}_d| f_c}{v} + \Theta_c, \ c = 1, 2, ..., 50. \tag{4}$$

Even if we have an identical $\vec{R}_d$ in the two data domains (i.e., the same subject and same movements), the sampled phase could still be very different when the antennas are deployed differently (i.e, giving a different $\vec{R}_s$). Consequently, different antenna deployment scenarios will have an impact on the RFID phase distortion, causing considerable divergence between the datasets sampled from different environments.

Unlike tag sensitivity variations, environment changes generate a specific type of phase distortion for all sampled phase data. Thus, the model variables in the deep learning network should be trained and optimized to combat such phase distortion. Given all kind of possible deployment environments, it is a big challenge to generate a well optimized deep learning model, which is generalizable for all environments.

IV. Meta-learning based Solution

In Meta-Pose, we propose meta-learning as an effective technique to initialize the variables based on trained tasks or data domains so that network could be effectively fine-tuned later for a new data domain [12]. The model-agnostic meta-learning algorithm (MAML) [13] has been proposed to pre-train the network, so the model could produce a satisfactory generalization performance. In addition, the Reptile learning algorithm [17] has also been proposed as a representative meta-learning algorithm, which works nearly as well as MAML while having a lower computational complexity. In this paper, we leverage Reptile to pre-train the model for initialization of model variables, and fine-tuning with a small amount of new data when applied to a new data domain.

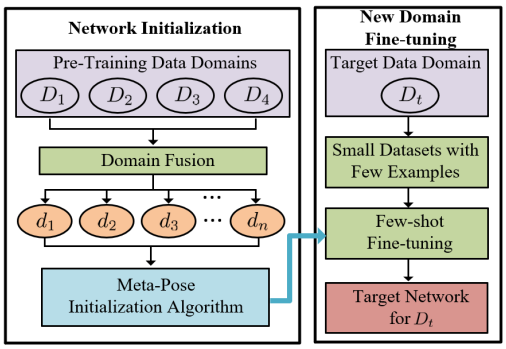

Fig. 3. Training framework of the proposed Meta-Pose system.

A. Meta-Pose Framework Overview

Figure 3 represents the brief structure of training framework of the proposed Meta-Pose system, which consists of network initialization and fine-tuning in a new domain. As shown in the figure, the deep learning model is trained with datasets from a few (e.g., four) known data domains, which are sampled when the subjects stand at four different positions. We notice that the performance of meta-learning can be improved by learning more learning tasks, but directly sampling a large amount of human pose data from numerous data domains is challenging and having a high cost. Thus, we propose a domain fusion algorithm to produce more data domains by mixing the data from the four available data domains. Then, we leverage the Reptile learning algorithm to recursively update the training variables of the network, so the variables will be well initialized. When transferring to the learning task in a new data domain, we only need to collect very few examples to fine-tune the generalized network.

B. Reptile based Network Initialization

The objective of network initialization is to determine the initial model variables, which could be adjusted for a new data domain with a few training steps. It means that the initial training variables H should be set close to any possible data domain D. The optimization problem for network initialization can be formulated as:

$$\min_{H} \quad \mathbb{E}_D[\Gamma(U_D^k(H))], \tag{5}$$

where the Γ denotes the loss function of the network, and $U_D^k(H)$ denotes the gradient decent operation that updates variables H for k times using data sampled from D, which is the Adam algorithm. According to the formulated problem, we propose the Meta-Pose initialization algorithm to generate the initial training variables H that facilitate fine-tuning, as presented in Algorithm 1. In the algorithm, we first fuse the four data domains (i.e., D_1, D_2, D_3, and D_4) into multiple fused data domains (i.e., $d_1, d_2, ..., d_n$). Each d_i contains 40 batches of data randomly sampled from D_1, D_2, D_3, and D_4.

To solve the optimization problem (5), we need to find the gradient of any fused data domain $\Delta\Gamma[U_{d_i}^k(H)]$, so the gradient decent algorithm can be applied to find H by recursive updating. With the Reptile learning algorithm [17], we first calculate $\Delta\Gamma[U_{d_i}^1(H)]$ for each inner loop iteration as:

$$\begin{aligned}\Delta\Gamma[U_{d_i}^1(H_{in})] =& U_{d_i}^1(H_{in}) - H_{in} \\ =& H'_{in} - H_{in},\end{aligned} \tag{6}$$

where H_{in} is the set of variables used in the inner loop. In the algorithm, denote the one step gradient $\Delta\Gamma[U_{d_i}^1(H_{in})]$ as W_j. The overall gradient after k iterations is calculated as:

$$\Delta\Gamma[U_{d_i}^k(H)] = \sum_{j=1}^{m} W_j.$$

$\Delta\Gamma[U_{d_i}^k(H)]$ is denoted as $\hat{W}_i$ for each data domain d_i. In the algorithm, we set $k = 8$ for effective training in each data domain. With gradient $\hat{W}_i$, we solve the problem by recursively training variable H in the outer loop iterations as:

$$H \leftarrow H + \epsilon\hat{W}_i, \tag{7}$$

where ϵ is the learning rate, which is set to 0.1 in the system. We repeat the updating process for $5,000$ times (i.e., setting $n = 5000$), so the final training result H_t could satisfy the requirement of the optimization problem (5). After initialization training, the network could be quickly fine-tuned with few shots of data sampled from a new data domain.

Algorithm 1: Meta-Pose Initialization Algorithm in the Meta-Pose System

Input: Sampled data sets from four data domains (denoted by D_1, D_2, D_3, and D_4);
Output: Optimally initialized variables H_t for the pretrained network.
Randomly initialize the training variable as H;
for $i = 1 : n$ **do**
  Generate d_i by randomly sampling from D_1, D_2, D_3, and D_4;
  Randomly sample k batches from d_i;
  $H_{in} \leftarrow H$;
  for $j = 1 : k$ **do**
    Update the variables in H_{in} with loss function Γ as: $H'_{in} = U_{d_i}^1(H_{in})$, $W_j = H'_{in} - H_{in}$, $H_{in} \leftarrow H'_{in}$;
  end
  Calculate the overall weight updates as: $\hat{W}_i = \sum_{j=1}^{k} W_j$;
  Update variables H as: $H \leftarrow H + \epsilon\hat{W}_i$;
end
Set $H_t \leftarrow H$;

C. Few-shot Fine-tuning

After appropriate initialization of H, the fine-tuning process only requires a small dataset from the new data domain. Since the training data are all data sequences, including RFID phase data and vision data [11], the data shots are defined specifically in the Meta-Pose system. We divide the data sequence into small segments during the training process, each consisting of

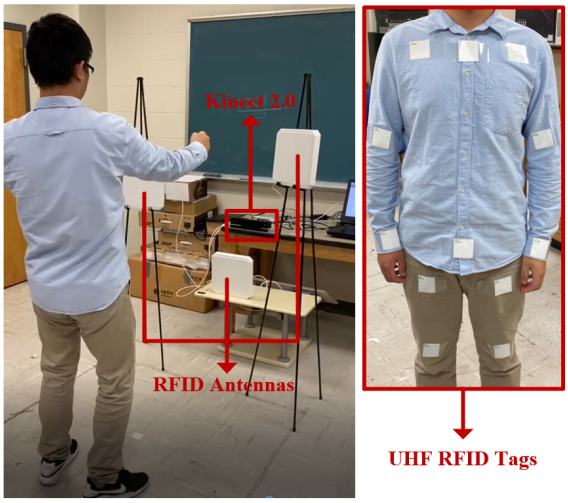

Fig. 4. Hardware configuration of the Meta-Pose system.

30 consecutive data samples sampled within a window of 6s. We consider one data batch as one shot in Meta-Pose, and less than 5 batches of data in the new data domain will be leveraged for fine-tuning. We also find that the type of movements also affects the fine-tuning performance and will discusses this in Section V-C. Due to the Reptile based initialization of training variables, the Meta-Pose system can quickly adapt to the new data domain with few-shot fine-tuning.

V. Implementation and Evaluation

A. Prototype System Implementation

To evaluate the performance of Meta-Pose, we develop a prototype system with an off-the-shelf Impinj R420 reader, which is configured with three S9028PCR polarized antennas, as shown in Fig. 4. The ALN-9634 (HIGG-3) RFID tags are used in Meta-Pose. The vision data, used for training supervision as well as ground truth for evaluating the precision of inference, is collected with an Xbox Kinect 2.0 device. As shown in the figure, we attach 12 RFID tags on the 12 joints of the subject, including the neck, pelvis, left hip, left knee, right hip, right knee, left shoulder, left elbow, left wrist, right shoulder, right elbow, and right wrist. With the three reader antennas placed at different height positions, every RFID tag can be interrogated by at least one of the antennas.

Environment adaption is validated using RFID data collected from eight different data domains, which are generated by the specific deployment of the subject and antennas as shown in Fig. 5. Seven data domains are sampled in the computer lab, and the eighth domain is sampled in an empty corridor. Among these domains, D_1 to D_4 are used for model pretraining, while D_5 to D_8 are considered as new data domains for evaluation. RFID phase data is collected when the subject stands in front of the antennas and performing specific activities repeatedly. Different types of activities are sampled in all the data domains, such as walking, body twisting, deep squatting, and single limb moving. Five subjects participate in the data sampling, including four males and one female.

B. Overall Performance Evaluation

To demonstrate the overall system performance, we use the 3D human skeleton data collected by Kinect 2.0 as ground truth. For each video frame, we calculate the mean error Ψ_{all} of all the 12 human joints as:

$$\Psi_{all} = \frac{1}{12}\sum_{n=1}^{12} ||\hat{T}_n - \dot{T}_n||, \tag{8}$$

where $\hat{T}_n$ represents the estimated 3D position of joint n, while $\dot{T}_n$ is the ground truth. $||\hat{T}_n - \dot{T}_n||$ is the Euclidean distance between the two 3D coordinates.

The overall performance (i.e., mean errors) of the fine-tuned network for all the eight data domains is presented in Fig. 6. Note that only the first four data domains are used in network pretraining, while the other four domains are used for testing. In addition, we also resent the accuracy of the pretrained network in the figure (i.e., without fine-tuning with additional data from the new data domain). As shown in the figure, the maximum error of the fine-tuned network is 4.83cm obtained in D_6, while the minimum error is 3.46cm obtained in D_8. The minimum pretraining error for the new data domain (i.e. D_5 to D_8) is 4.91cm in D_8, which is higher than that of all the pretrained domains (i.e. D_1 to D_4). The higher pretained errors imply the large divergence between the known and new data domains. However, with few-shot fine-tuning, the mean error for all the four new data domains is 3.98cm, which is very similar to that of the pretrained data domains. The considerable error reduction in D_4 to D_8 is due to the Meta-Pose initialization algorithm. With well optimized training variables, the network can be effectively fine-tuned for new data domains. Compared to the height of the subject and range of motions, the 3D human pose estimation errors are small and negligible. These results demonstrate the high adaptability of the Meta-Pose system.

C. Fine-tuning Evaluation

For most effective fine-tuning, we also conduct experiments to investigate the impact of numbers of shots and types of activities. Fig. 7 illustrates the accuracy of pose tracking in the four new data domains, which are fine-tuned with different numbers of data shots from 1 to 5. As defined earlier, one-shot of data in Meta-Pose is defined as a consecutive data sequence within a window of 6 seconds. It can be seen that, after 5-shot fine-tuning, the minimum error 3.49cm is achieved in D_8, while the error in D_6 is the highest (i.e., 4.68cm). In addition, although the final estimation accuracy is different for the four data domains, the performance of fine-tuning is generally improved by more data shots. However, as the figure shows, the improvement becomes not obvious beyond four shots of data. Thus, 4-shot fine-tuning is sufficient when the Meta-Pose system is transferred to a new environment.

D. Comparison with a Baseline Scheme

We also conducted a comparison study using the recent RFID based pose tracking system RFID-Pose as a baseline scheme [10]. We leverage the same training data to perform

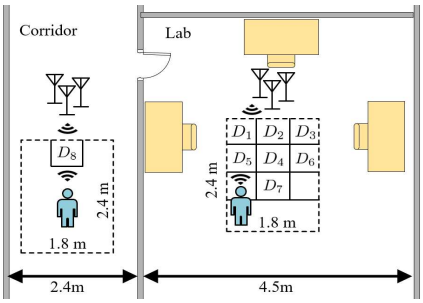

Fig. 5. Illustration of the data domains used in the Meta-Pose experiments.

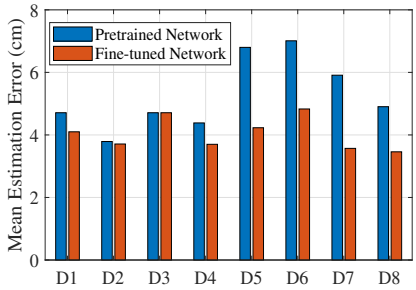

Fig. 6. Overall performance in terms of mean estimation error in the eight different data domains.

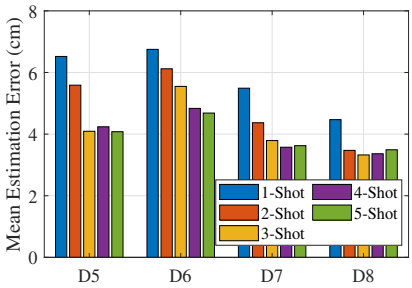

Fig. 7. Fine-tuning performance of different new data domains with different shots of new data.

TABLE I
PERFORMANCE COMPARISON AFTER FINE-TUNING

Domain Index	*RFID-Pose*	*Meta-Pose*
D_5	6.72cm	3.72cm
D_6	7.62cm	4.32cm
D_7	5.46cm	3.51cm
D_8	4.62cm	4.11cm
D_{all}	6.27cm	3.97cm

4-shot fine-tuning for each unknown data domain. The pose tracking errors are presented in Table I. As the table shows, the mean error of RFID-Pose for all the new data domains is 6.27cm, while that for Meta-Pose is only 3.97cm. We find that the RFID-Pose error is also reduced by fine-tuning, but the estimation error for the new data domain is still quite high. A larger datasets sampled in the new environments are needed to achieve a satisfactory fine-tuning performance, which considerably increases the training data collection effort and cost. In contrast, the error of Meta-Pose has been effectively reduced by few-shot fune-tuning, because the meta-learning-based algorithm has suitably initialized the model variables based on the known data domains. Meta-Pose is able to quickly optimize its training variables for the untrained data domain with a few steps of gradient descent. Through these experiments, we demonstrate that Meta-Pose can better adapt to unknown environments compared with the baseline scheme. Thus it can be easily deployed in practice.

VI. CONCLUSIONS

In this paper, we proposed an RFID based realtime 3D pose tracking system named Meta-Pose that is environment-adaptive. A novel Meta-Pose initialization algorithm was proposed to pretrain the network with several known data domains, and few-shot fine-tuning was then utilized to adapt to unknown data domains. The prototype Meta-Pose system was constructed with commodity RFID reader and tags. Extensive experiments were conducted with ground truth provided by Kinect vision data. The high adaptability to new environments was demonstrated by our experimental results and a comparison study with a state-of-the-art baseline scheme.

ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation under Grants ECCS-1923163 and CNS-2107190.

REFERENCES

[1] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proc. IEEE CVPR 2017*, Honolulu, HI, July 2017, pp. 7291–7299.
[2] M. Andriluka, S. Roth, and B. Schiele, "Monocular 3D pose estimation and tracking by detection," in *Proc. IEEE CVPR 2010*, San Francisco, CA, June 2010, pp. 623–630.
[3] Tom's Guide, "Millions of wireless security cameras are at risk of being hacked: What to do," 2020 (accessed Aug. 28, 2020). [Online]. Available: https://www.tomsguide.com/news/hackable-security-cameras
[4] M. A. A. da Cruz, J. J. P. C. Rodrigues, P. Lorenz, V. Korotaev, and V. H. C. de Albuquerque, "In.iot—a new middleware for internet of things," *IEEE Internet of Things Journal*, pp. 1–10, Dec. 2020.
[5] M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, "Through-wall human pose estimation using radio signals," in *Proc. IEEE CVPR 2018*, Salt Lake City, UT, June 2018, pp. 7356–7365.
[6] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10 032–10 044, Sept. 2020.
[7] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang, "Person-in-WiFi: Fine-grained person perception using WiFi," in *Proc. IEEE ICCV 2019*, Seoul, Republic of Korea, Oct. 2019, pp. 5452–5461.
[8] W. Jiang, H. Xue, C. Miao, S. Wang, S. Lin, C. Tian, S. Murali, H. Hu, Z. Sun, and L. Su, "Towards 3D human pose construction using WiFi," in *Proc. ACM MobiCom'20*, London, UK, Sept. 2020, pp. 1–14.
[9] J. Zhang, S. Periaswamy, S. Mao, and J. Patton, "Standards for passive UHF RFID," *ACM GetMobile*, vol. 23, no. 3, pp. 10–15, Sept. 2019.
[10] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, in press. DOI: 10.1109/TR.2020.3030952.
[11] ——, "Subject-adaptive skeleton tracking with RFID," in *Proc. IEEE MSN 2020*, Tokyo, Japan, Dec. 2020.
[12] J. Vanschoren, "Meta-learning: A survey," arXiv preprint arXiv:1810.03548, Oct. 2018. [Online]. Available: https://arxiv.org/abs/1810.03548
[13] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. ICML 2017*, Sydney, Australia, Aug. 2017, pp. 1126–1135.
[14] M. Lenehan, "Application note - low level user data support," Feb. 2019. [Online]. Available: https://support.impinj.com/hc/en-us/articles/202755318-Application-Note-Low-Level-User-Data-Support
[15] R. Villegas, J. Yang, D. Ceylan, and H. Lee, "Neural kinematic networks for unsupervised motion retargetting," in *Proc. IEEE CVPR 2018*, Salt Lake City, UT, June 2018, pp. 8639–8648.
[16] L. Ukkonen and L. Sydanheimo, "Threshold power-based radiation pattern measurement of passive UHF RFID tags," *PIERS Online*, vol. 6, no. 6, pp. 523–526, 2010.
[17] A. Nichol, J. Achiam, and J. Schulman, "On first-order meta-learning algorithms," arXiv preprint arXiv:1803.02999, Oct. 2018. [Online]. Available: https://arxiv.org/abs/1803.02999