Demo Abstract: Environment-adaptive 3D Human Pose 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, lzw0039}@auburn.edu, xuyu.wang@csus.edu, smao@ieee.org

***Abstract*—RF-based human pose estimation has attracted increasing interest in recent years. Compared with vision-based approaches, RF-based techniques can better protect user's privacy and are robust to lighting and non-line-of-sight conditions. However, due to complicated indoor propagation environments, most of the RF-based sensing approaches are sensitive to the deployment environment and hard to adapt to new environments. In this demo, we present a meta-learning-based approach to address the environment adaptation problem and design an environment-adaptive Radio-Frequency Identification (RFID) based 3D human pose tracking system. The system utilizes commodity RFID tags to estimate 3D human pose and leverage meta-learning algorithms to improve the environment adaptability. Experiments conducted in various environments demonstrate the high pose estimation performance and adaptability to environments.**

***Index Terms*—3D human pose estimation, meta-learning, MAML, Reptile, RFID sensing.**

I. Introduction

Human posture estimation and tracking are useful for a variety of applications, such as human-computer interaction, video surveillance, and somatosensory gaming. Nevertheless, the video data collected for pose monitoring could be intercepted by attackers, thus raising security and privacy concerns. Radio Frequency (RF) based approaches have been proposed to address the privacy issue, but they also suffer from poor generalization when applying a trained system to different environments. This is because the well-trained deep learning model used in such systems is usually hard to apply to the untrained data domains sampled in new RF environments. To address the environment adaptation issue, various approaches, such as data augmentation, domain adversarial networks, and transfer learning, have been proposed recently.

In this demo, we leverage a meta-learning strategy to address the environment/domain adaptation problem and present *Meta-Pose* [1], [2], which is an environment-adaptive, RFID-based 3D human pose tracking system. In Meta-Pose, RFID tags are attached to the human body, as in our previous work RFID-Pose [3], so that the motion of human joints could be estimated from the RFID phase data collected by the reader. Unlike previous works, Meta-Pose is pretrained by meta-learning algorithms with training data sampled from a small number of known environments. The meta-learning algorithms are adopted to achieve an optimized network initialization, so the system is able to adapt to an untrained environment with few-shot fine-tuning. The high environment adaptability of the prototype system is demonstrated by our experiments conducted in various indoor environments.

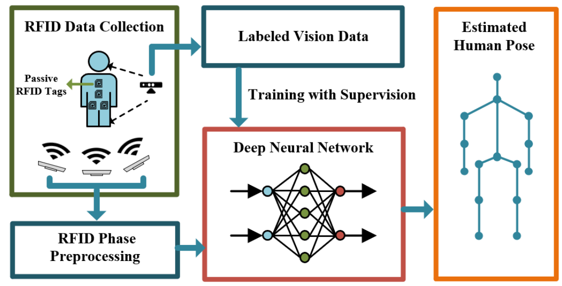

Fig. 1. Overview of the architecture of the proposed Meta-Pose system.

II. System Design

A. System Architecture Overview

The Meta-Pose system is proposed to perform 3D human pose tracking with phase data collected from the RFID tags attached on the joints of the subject. Fig. 1 presents an overview of the Meta-Pose system architecture, which is composed of three key modules, including RFID phase sampling, phase data preprocessing, and a multi-model deep neural network.

1) RFID Data Collection and Signal Preprocessing: As the prior RFID pose tracking approach [3], 3D human skeleton is estimated from the phase data measured through RFID communications with the Low Level Reader Protocol (LLRP). The phase variation of two consecutively sampled phases can alleviate the interference of the phase offset caused by channel hopping. Such phase variation data effectively captures over time the variations in the tag-antenna distance, which is an effective indicator of the subject's joint movements.

2) Vision-aided Deep Neural Network: The translation from phase variation data to 3D human pose is achieved by a multi-modal vision-aided deep neural network. The input to the network is the preprocessed phase variation data, and training is accomplished by the supervision of the 3D human pose generated by Kinect. A recurrent autoencoder is implemented in the network to extract the movement feature from RFID data and translate it to the movement of human limbs [3]. The mapping from the autoencoder output to the 3D coordinates of human skeleton is accomplished in the Forward Kinematics layer, which is a classic motion generator used in 3D animation and robotics [4]. The training goal is to

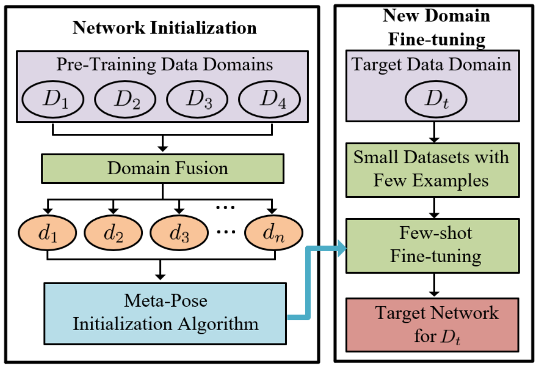

Fig. 2. Illustrate the training procedure of Meta-Pose.

minimize the error between the estimated human pose and the corresponding label (i.e., vision data), so the network should perform effective translation from RFID data to human pose.

B. Meta-learning for Environment Adaptation

The purpose of meta-learning is to better initialize the network parameters in pretraining, which can then be further improved when applied to a new data domain with only a few training examples. Therefore the use of meta-learning helps to improve the adaptability of the model and overcome the barrier of deploying the system in various environments. The meta-learning initialization structure is briefly illustrated in Fig. 2. The deep learning model is initially pretrained with the datasets sampled from four known data domains, each with a different subject location, antenna placement, and propagation environment. Two representative meta-learning initialization algorithms, MAML [5] and Reptile [6], are utilized for the pretraining. Once the network is properly initialized, only a few new training data will be required from a new data domain for fine-tuning the model.

III. Implementation and Evaluation

Extensive experiments in various RF environments are conducted to assess the system performance. The prototype system is created with a commodity Impinj R420 reader equipped with three polarized antennas S9028PCR. As Fig. 3 shows, 12 ALN-9634 (HIGG-3) RFID tags are attached to the clothes of the subject. The vision data used for vision-aided training and system evaluation is sampled by an Xbox Kinect 2.0 device. Eight different RF environments (named D_1 to D_8) are evaluated with different antenna deployments and locations, where D_1 to D_4 are used for model pretraining, and D_5 to D_8 are new data domains for evaluating model adaptability.

Figure 4 illustrates the Cumulative Distribution Functions (CDF) of the estimation error between vision ground truth and estimated 3D pose, which are respectively achieved by the proposed Meta-Pose system and the traditional pose estimation scheme RFID-Pose [3]. The estimation errors are obtained after four-shot fine-tuning for all the four untrained data domains (i.e., D_5 to D_8). As the figure shows, the median estimation error of the baseline is 6.87 cm, whereas the Meta-Pose achieves an 3.94 cm median estimation error. The

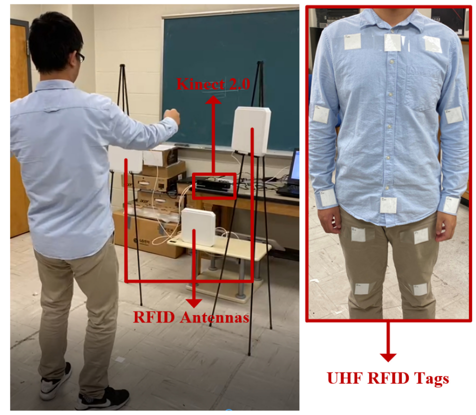

Fig. 3. Experiment set-up of the Meta-Pose system.

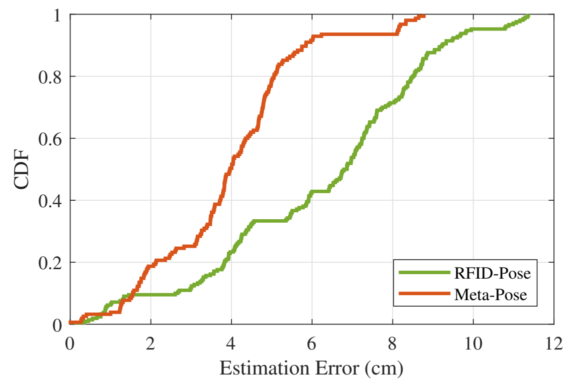

Fig. 4. The CDFs of four-shot fine-tuning achieved by Meta-Pose [2] and the baseline RFID-Pose [3].

experiment results demonstrate Meta-Pose's stronger ability to adapt to new environments than the baseline scheme.

Acknowledgments

This work is supported in part by the NSF under Grants ECCS-1923163 and CNS-2107190, and through the Wireless Engineering Research and Education Center (WEREC) at Auburn University.

References

[1] C. Yang, L. Wang, X. Wang, and S. Mao, "Meta-Pose: Environment-adaptive human skeleton tracking with RFID," in *Proc. IEEE GLOBECOM 2021*, Madrid, Spain, Dec. 2021, pp. 1–6.

[2] C. Yang, L. Wang, X. Wang, and S. Mao, "Environment adaptive RFID based 3D human pose tracking with a meta-learning approach," *IEEE Journal of Radio Frequency Identification*, in press. DOI: 10.1109/JRFID.2022.3140256.

[3] C. Yang, X. Wang, and S. Mao, "RFID-Pose: Vision-aided 3D human pose estimation with RFID," *IEEE Transactions on Reliability*, vol. 70, no. 3, pp. 1218–1231, Sept. 2021.

[4] 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.

[5] 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.

[6] 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